

Luca Fanelli, Paolo Paruolo

# Exchange rates, prices and their speed of adjustment

2006/7



UNIVERSITÀ DELL'INSUBRIA  
FACOLTÀ DI ECONOMIA

<http://eco.uninsubria.it>

I punti di vista espressi nei quaderni della Facoltà di Economia riflettono unicamente le opinioni degli autori, e non rispecchiano necessariamente quelli della Facoltà di Economia dell'Università dell'Insubria.

The views expressed in the working papers reflect the opinions of the authors only, and not necessarily the ones of the Economics Faculty of the University of Insubria.

© Copyright Luca Fanelli, Paolo Paruolo  
Printed in Italy in September 2006  
Università degli Studi dell'Insubria  
Via Monte Generoso 71, 21100 Varese, Italy

All rights reserved. No part of this paper may be reproduced in any form without permission from the Authors.

# Exchange rates, prices and their speed of adjustment\*

Luca Fanelli<sup>†</sup>, Paolo Paruolo<sup>‡</sup>

September 13, 2006

## Abstract

This paper addresses the problem of measuring the speed of adjustment of exchange rates and relative prices to purchasing power parity (PPP), in the multivariate context of Vector Autoregressive Processes (VAR). We consider the speed of adjustment of one variable  $y$  in response to another variable  $x$ , where  $x, y$  belong to the VAR. We propose a multivariate measure defined as the forecasting horizon for which the cumulated interim multiplier of  $x$  on  $y$  surpasses a given fraction  $p$  of the corresponding total multiplier. This measure of speed for  $p = \frac{1}{2}$  coincides with the usual concept of half-life when restricted to univariate processes. We emphasize the importance to separate the concepts of long run effect size and its speed of adjustment, where the latter is unambiguously defined only when the long run effect is non-zero. We discuss likelihood-based point estimators and confidence sets for this notion of half-life, and reconsider evidence on adjustment to PPP in monthly post-Bretton Woods data for five major industrialized countries against the U.S. dollar. Results show that nominal exchange rates buffer the entire adjustment to PPP disequilibrium, whereas relative prices do not adjust either in the short or the long run to PPP deviations. Concluding in such a situation that prices adjust faster than exchange rates is a matter of how one interprets the absence of short run and long run effects.

**Keywords:** Half-life, purchasing power parity, impact factors, speed of adjustment, vector equilibrium correction, upcrossing and downcrossing.

**J.E.L. Classification:** C32, C52, F31.

---

\*Paper presented at ‘Frontiers of Econometric Forecasting - Second workshop on dynamic econometric in memory of Carlo Giannini’, June 9, 2006, Brescia Italy. We thank Rocco Mosconi and Riccardo ‘Jack’ Lucchetti for useful comments. Partial financial support is gratefully acknowledged from MIUR ex-60% grants (first author), MIUR grants Cofin2004 and University of Insubria FAR 2005-2006 (second author).

<sup>†</sup>Corresponding author. University of Bologna, Department of Statistical Sciences, via Belle Arti 41, I-40126 Bologna, Italy. E-mail: fanelli@stat.unibo.it.

<sup>‡</sup>Department of Economics, Via Monte Generoso 71, I-21100 Varese, Italy. Email: paolo.paruolo@uninsubria.it.

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Motivating example</b>	<b>3</b>
<b>3</b>	<b>Definitions</b>	<b>6</b>
3.1	VAR . . . . .	6
3.2	Impact factors and interim multipliers . . . . .	7
3.3	Half-life . . . . .	9
3.4	Illustration . . . . .	11
<b>4</b>	<b>Inference</b>	<b>12</b>
4.1	Inference on long run effects . . . . .	12
4.2	Inference on the half-life . . . . .	13
<b>5</b>	<b>Empirical analysis</b>	<b>14</b>
5.1	Data . . . . .	14
5.2	Specification analysis . . . . .	15
5.3	Empirical findings . . . . .	17
<b>6</b>	<b>Conclusions</b>	<b>18</b>
	<b>References</b>	<b>19</b>
<b>A</b>	<b>Appendix</b>	<b>21</b>
A.1	Eigenvalues and half-life . . . . .	21
A.2	Nonstable eigenvalues . . . . .	22
A.3	Perturbation of eigenvalues . . . . .	23
A.4	Grid on the ellipsoid boundary . . . . .	24
	<b>Figures and Tables</b>	<b>26</b>

# 1 Introduction

The speed of adjustment of prices and exchange rates to deviations from equilibrium is receiving increasing attention in the empirical debate on purchasing power parity (PPP). According to the traditional sluggish-price explanation, prices should not adjust to equilibrium faster than nominal exchange rates, consistently with the rational expectations sticky-price paradigm of Dornbusch (1976) and Mussa (1982), see also Kim (2005). Some authors have recently challenged this view, arguing that the empirical evidence over the post-Bretton Woods period is supportive of quicker price adjustment, see Engel and Morley (2001) and Cheung et al. (2004).

The notion of half-life as a measure of speed of adjustment to the PPP has been used extensively. Despite empirical questions are most naturally posed in terms of the effects of one variable  $x$  on another  $y$ , the concept of half-life most often employed in the literature has been a univariate one, see e.g. Cheung and Lai (2000), Kilian and Zha (2002) and Rossi (2005), *inter alia*. Unfortunately, univariate methods do not allow to disentangle the contribution of nominal exchange rates and prices on real exchange rates dynamics, hence on the half-life of PPP deviations. Although more recent papers attempt to measure the speed of adjustment via impulse responses in Vector Autoregressive (VAR) models, see Cheung et al. (2004) and Crowder (2004), a formal definition of half-life in this multivariate context appears to be lacking.

The present paper provides a formal definition of half-life – as well as of more general measures of speed of adjustment – in the multivariate framework of (possibly integrated and cointegrated) VARs. The proposed definition is then applied to investigate the speed of PPP adjustment of prices and exchange rates for the major industrialized countries over the post-Bretton Woods period.

The definition of half-life we introduce is based on the companion form associated with the Vector Error Correction (VEC) representation of systems integrated of order one,  $I(1)$ . It is hence based on the whole system dynamics, and provides a unified framework for measuring speeds of adjustment of different  $(y, x)$  pairs of variables. Despite the interpretation of half-lives described in this paper is most natural in the context of  $I(1)$  systems, our definitions equally apply to stationary VARs.

In line with Andrews and Chen (1994), we emphasize that an unambiguous definition of half-life is possible only if  $x$  has a non-zero long run effect on  $y$ . This observation suggests to distinguish between the ‘size’ of the long run effect of  $x$  on  $y$ , and the ‘speed’ at which this effect is accomplished, if the effect size is non-zero. It is found that the present definition of long run effect size coincides with the notion of ‘impact factors’ (IF) defined in Omtzigt and Paruolo (2005), which is in turn strictly related to the ‘total multiplier’ equal to the sum of all the moving average coefficient matrices of a VAR.

These long run effects are contrasted in this paper with short run effects that can be deduced for any finite forecasting horizon. The two concepts are found to be different, in the sense that for a single pair of variables  $(x, y)$ , a zero or non-zero long run effect does not imply nor is implied by a zero or non-zero short run effect. Omtzigt and Paruolo (2005) give a detailed account on how to test for zero long run effect size using likelihood-based inference, after fixing the cointegration rank. In the present paper we extend their analysis by deriving point estimators and confidence sets for half-lives, which are based on the likelihood estimation of the VEC.

Our approach has several advantages. First, it allows to separate inference on the speed of adjustment in two steps. In the first step we test if the size of the long

run effect (i.e. the IF) of  $x$  on  $y$  is zero. If this hypothesis is rejected we proceed to estimate the corresponding half-life and we construct an asymptotic confidence set for it. If the long run effect size is instead insignificant, we suggest not to measure any speed of adjustment, given that there is not enough sample information to establish if there is any long run effect at all.

Secondly, this approach allows to investigate several aspects of PPP adjustment in a consistent and comprehensive framework, without the necessity to resort to any structural interpretation of VAR disturbances. In a bivariate I(1) system with log exchange rates,  $e$ , and log relative prices,  $p$ , one can first test the PPP implication that the cointegration rank,  $r$ , is equal to 1. This test can be interpreted as a verification that one equilibrium exists in the system as predicted by PPP. If  $r = 1$ , one can then test if this equilibrium relation indeed corresponds to the real exchange rate  $q := e - p$ , again as predicted by PPP. In the same system one can then proceed to measure the speed of adjustment, using the present approach. In particular, one can compute the speed of adjustment of PPP deviations in response to respectively nominal exchange rate and price variations. Likewise, one can measure the speed of adjustment of nominal exchange rates and prices to PPP deviations. A single misspecification analysis performed on the unrestricted VAR prior to any of these tests of the PPP and to the calculation of the half-lives can guard against misspecified inferences.

We apply these concepts to the analysis of the speed of PPP adjustment of prices and exchange rates for five industrialized countries over the post-Bretton Woods period. Differently from Crowder (2004), who also uses VECs to investigate the convergence of nominal exchange rates and prices to PPP, we do not appeal to any structural identification of shocks. As in Klein et al. (1991), we argue that PPP adjustment can be investigated by considering both periods before and after the Plaza Agreement of 1985. Focusing on the most recent period, there is evidence of mean-reversion of real exchange rates, with point estimates of the half-life of PPP deviations below one year, though upper bounds of confidence sets do not allow to dismiss the 3 to 5 years consensus view.

Results over the entire 1973-1998 period are more consistent with the recent literature, see e.g. Elliott and Pesavento (2006). However, in this sample, even fixing the cointegration rank to one, the highest eigenvalue of the companion matrix of the estimated VECs is very close to one; questioning the possibility of tenable inference on half-lives of PPP deviations. These empirical results are hence more fragile, and we focus attention to the 1985-1998 subsample.

A common finding for all countries is the absence of any short run and long run adjustment of relative prices to PPP, whereas all the adjustment seems to be accomplished by nominal exchange rates. This evidence can be compared with recent papers on PPP adjustment which use multivariate models. For instance, Cheung et al. (2004) use VECs and Pesaran and Shin's (1996, 1998) generalized impulse response (GIR) analysis and conclude that relative prices adjust faster than nominal exchange rates for the major industrialized countries over the post-Bretton Woods period. Their conclusion is the same as Engel and Morley (2001), who use unobserved component models. Both papers invoke new economic explanations for their interpretation that nominal prices adjust relatively quickly, that nominal and real exchange rates are highly volatile, and that nominal exchange rates converge very slowly to the PPP equilibrium. They argue that the central PPP puzzle consists in how to explain why nominal exchange rates converge so slowly.

The interpretation of Engel and Morley (2001) and Cheung et al. (2004) seems to contrast with the findings of other researchers more in line with the sluggish-price tradition, e.g. Kim (2005) and Cecchetti et al. (2002). The analysis in Kim (2005) is based on a modified version of Mussa's (1982) model with traded and nontraded goods, and reports faster adjustment speed of exchange rates to PPP compared to the previous literature. Likewise, Cecchetti et al. (2002) find that PPP convergence between cities within the U.S. is much slower than the one observed at the international level, arguing therefore that nominal exchange rates facilitate the speed of adjustment.

Our empirical results show that relative prices do not adjust at all. Interpreting this evidence along the lines of Cheung et al. (2004), as indicating that prices adjust to PPP equilibrium faster than nominal exchange rates, is a matter of definition. Consistently with our distinction between long run effect size and its speed introduced previously, we interpret the fact that relative prices do not adjust as a situation where one cannot measure the speed of adjustment, given that no adjustment takes place. Hence our empirical evidence appears in line with evidence in the above recent papers, although our interpretation differs, thanks to the separation between the concepts of long run effect and speed of adjustment.

The rest of the paper is organized as follows. Section 2 introduces the main issues of the paper by a motivating example. Section 3 discusses the definition of half-life within the context of I(1) VARs. Section 4 summarizes results concerning likelihood-based inference for half-lives. Section 5 investigates the half-life of PPP adjustment for nominal exchange rate and relative prices for 5 major industrialized countries, and Section 6 contains some concluding remarks. Technical details and proofs are reported in the Appendix.

## 2 Motivating example

In this section we present a motivating simple example; this example is included to illustrate the main concepts discussed in the paper, and it does not strive either to be general or to resemble reality. We consider bivariate Data Generating Processes (DGP), for which we discuss long versus short run adjustment, in preparation for the general definitions, which are introduced in Section 3.

Define a bivariate system  $X_t := (X_{1t} : X_{2t})'$ , where the two scalars  $X_{1t}$  and  $X_{2t}$  may represent e.g. log exchange rates and log relative prices, as in the empirical application. Assume that the system  $X_t$  is generated by:

$$\Delta X_{1t} = -\frac{1}{2}(X_{1t-1} - X_{2t-1}) + \varepsilon_{1t} \quad (1)$$

$$\Delta X_{2t} = \gamma \Delta X_{1t-1} + \varepsilon_{2t} \quad (2)$$

where  $X_0 = (X_{10} : X_{20})'$  is fixed,  $\varepsilon_t = (\varepsilon_{1t} : \varepsilon_{2t})'$  is i.i.d. Gaussian with positive definite covariance matrix  $I_2$ , the identity matrix of order 2. Here  $\Delta := 1 - L$  is the difference operator and  $L$  is the lag operator. We label the DPG with  $\gamma = 0$  as DGP<sub>1</sub>, and the one with  $\gamma = \frac{1}{4}$  as DGP<sub>2</sub>.

Application of the conditions in Granger's Representation Theorem (see e.g. Johansen 1996 Theorem 4.2) to these DGPs shows that  $X_t$  is I(1) in both cases, with cointegration rank  $r = 1$  and cointegrating vector  $\beta = (1 : -1)'$ . Note that the representation given in (1)-(2) is a VEC form. The process  $q_t := X_{1t} - X_{2t}$  measures

the deviations from the long run equilibrium, hereafter the disequilibrium term. In the case of PPP,  $q_t$  measures the log real exchange rate.

The main question we address in this paper is the following: ‘which variable in a VEC like (1)-(2) adjusts faster to long run equilibrium?’ We try to address this question by decomposing it into the following two questions. Q<sub>1</sub>: ‘does  $y$  adjust to variations in  $x$ ?’ and Q<sub>2</sub>: ‘if  $y$  adjusts to variations in  $x$ , what is the speed of adjustment?’.

We first consider Q<sub>1</sub>. Candidate response variables  $y$  are in this case  $X_{1t}$ ,  $X_{2t}$  as well as  $q_t$ . The cause variable  $x$  may be taken to be  $q_t$ , given that one is interested in adjustment to equilibrium, and  $q_t$  represents deviations from equilibrium. However, one may wish to consider effects on  $y$  of any variable that causes  $q_t$ , like both  $X_{1t}$  and  $X_{2t}$ . This shows that there are many possible interesting choices for  $x$  and  $y$ .

We next introduce short run and long run effects. The error correction coefficient  $-\frac{1}{2}$  in the VEC represents the response of  $y_{t+1} := \Delta X_{1t+1}$  to  $x_t := q_t$  for one-step ahead forecasts. Formally we take

$$G_{y,x}(i) := \frac{\partial E_t y_{t+i}}{\partial x_t}$$

as a definition of the short run effect of  $x_t$  on  $y_{t+i}$ , where  $E_t$  indicates conditional expectations given the information set generated by  $X_{t-s}$ ,  $s \geq 0$ . Note that  $E_t y_{t+i}$  is the point predictor of  $y_{t+i}$  based on information up to and including time  $t$ , so that the effect of  $x_t$  on  $y_{t+i}$  is defined as the multiplier of  $x_t$  in the point forecast of  $y_{t+i}$ . We call this effect a short run effect because it converges to 0 as the forecast horizon increases to  $\infty$ , see Section 3.

The corresponding long run effect of  $x$  on  $y$  is defined instead as the cumulation of all the short run effects of  $x_t$  on  $y_{t+i}$  for all forecast horizons  $i = 1, \dots, \infty$ . In the present case, because  $y_{t+i} = \Delta X_{1t+i}$  measures a growth rate, this cumulated effect coincides with the effect of  $x_t$  on  $X_{1\infty} - X_{1t}$ , i.e. on the total ‘displacement’ of  $X_{1t+i}$  caused by  $x_t$ . Hence the long run effect, indicated by  $F_{y,x}$ , corresponds to an effect on the levels of  $X_{1t}$  caused by  $x_t := q_t$ . Using the formulas given in Section 3 below, one can easily compute  $F_{y,x}$  for the present choice of  $y$  and  $x$ , which equals  $F_{\Delta X_1, q} = -1$  for DGP<sub>1</sub> and  $F_{\Delta X_1, q} = -\frac{4}{3}$  for DGP<sub>2</sub>. One hence sees that the short run and long run effects differ in size, and that the long run effect of  $x$  on  $y$  may well depend on other coefficients in the dynamics that are apparently unrelated to  $x$  or  $y$ ; in this case it is seen that the long run effect of  $q$  on  $\Delta X_1$  depends on the value taken by  $\gamma$ .

Let us now compare these short run and long run effects of  $x_t = q_t$  on  $X_{1t}$  with the ones on  $X_{2t}$ ; i.e. let  $x_t := q_t$  and  $y_{t+i} := \Delta X_{2t+i}$ . The short run effect for  $i = 1$  is seen to be equal to 0 both for DGP<sub>1</sub> and DGP<sub>2</sub>. Again using the formulas given in Section 3 below, one finds  $F_{\Delta X_2, q} = 0$  for DGP<sub>1</sub> and  $F_{\Delta X_2, q} = -\frac{1}{3}$  for DGP<sub>2</sub>. Hence the long run effect of  $x$  on  $y$  may be 0 in the short run and zero or non-zero in the long run.

This example demonstrates how short and long run effects are different concepts in the sense that any or both the short run and long run effects may be zero or non-zero. In the following we will concentrate on the long run effects. We say that  $y$  adjusts to variations in  $x$  if the long run effect of  $x$  on  $y$  is non-zero.

We now consider question Q<sub>2</sub>: ‘if  $y$  adjusts to variations in  $x$ , what is the speed of adjustment?’ In order to measure speed of adjustment we consider the ratio  $f_{y,x}(\ell)$



of the cumulated short run effect up to some horizon  $\ell$  to the long run effect, i.e.

$$f_{y,x}(\ell) := \frac{\sum_{i=1}^{\ell} G_{y,x}(i)}{F_{y,x}}, \quad \ell = 1, 2, \dots$$

This ratio is defined only when  $y$  adjusts to  $x$ , i.e. when there is a non-zero long run effect of  $x$  on  $y$ ,  $F_{y,x} \neq 0$ .  $f_{y,x}(\ell)$  is used as a basis for the definition of half-life, as explained below.

Observe that by definition  $f_{y,x}(\ell)$  converges to 1 when  $\ell \rightarrow \infty$ , so that there is a first time,  $u_1$  say, when  $f_{y,x}(\ell)$  surpasses a given fraction  $p$ , where  $0 < p < 1$ , and usually we take  $p = \frac{1}{2}$ . We name  $u_1$  the first time of upcrossing of level  $p$ . Given that  $f_{y,x}(\ell)$  is not necessarily monotonic, there may be other horizons  $\ell > u_1$  where  $f_{y,x}(\ell)$  falls below  $p$ . In this case we define  $u_2$ , the second time of upcrossing, in a similar fashion. This process is repeated until the short run effects converge to 0 and hence  $f_{y,x}(\ell)$  approaches 1. We let  $u_{\max}$  denote the last time of upcrossing.

One may wish to take as definition of half-life the first time of upcrossing  $u_1$ . This is the minimal time taken by the system in order to reach at least 50% of the long run effect of  $x$  on  $y$ . Others may prefer to take  $u_{\max}$  as the definition of half-life; for instance, Kilian and Zha (2002) suggest such an approach. This represents the time the system takes to fluctuate below and around the 50% of the long run effect of  $x$  on  $y$ . In other words  $u_{\max}$  is the time after which the fraction of long run effect of  $x$  on  $y$  that has materialized is never below 50%. Other economists may prefer to make the median time of upcrossing as the proper definition of half-life. Obviously, there are many options. We define as half-life of the long run effect of  $x$  on  $y$  as *any* of these times of upcrossing, depending on the specific goals of the research exercise.

Some of the possible shapes of  $f_{y,x}(\ell)$  are illustrated in Figures 1 and 2, using DGP<sub>1</sub> and DGP<sub>2</sub>, as well as two additional DGPs, labelled DGP<sub>3</sub> and DGP<sub>4</sub>. They are defined as bivariate VAR(1) processes  $X_t := (X_{1t} : X_{2t})'$  with autoregressive matrix

$$\rho \begin{pmatrix} \cos \psi & \sin \psi \\ -\sin \psi & \cos \psi \end{pmatrix}.$$

In particular, DGP<sub>3</sub> corresponds to  $\rho = 0.7$ ,  $\psi = 0.25\pi$ , while DGP<sub>4</sub> to  $\rho = -0.7$ ,  $\psi = 0$ . In Figures 1 and 2 the subscripts  $y$  and  $x$  of  $f_{y,x}(\ell)$  are substituted by the corresponding position of variables  $y$  and  $x$  in the state vector. For DGP<sub>1</sub> and DGP<sub>2</sub>, '1' corresponds to  $\Delta X_{1t}$ , '2' to  $\Delta X_{2t}$ , and '3' to  $q_t := X_{1t} - X_{2t}$ . For DGP<sub>3</sub> and DGP<sub>4</sub>; '1' corresponds to  $X_{1t}$ , '2' to  $X_{2t}$ .

The plots in Figure 1 show how  $f_{y,x}(\ell)$  may display monotonic and non-monotonic behavior. These may induce one or many upcrossings of the level  $p$ . Note that varying  $p$  one obtains different level sets that give different information on  $f_{y,x}(\ell)$ . The present definition hence generates a family of indicators for the speed of adjustment, which includes the one of half-life as a special case.

Before concluding this section we illustrate how the present definitions relate to the debate on the speed of adjustment of random walks. A part of the literature has stated that a random walk adjusts instantaneously, see Morley (2006), footnote 2 and pp. 14-15, and Engel and Morley (2001), pp. 20-21. This case is represented by  $y_t := \Delta X_{2t}$ ,  $x_t := \Delta X_{2t}$  in (2) for DGP<sub>1</sub>. We first ask Q<sub>1</sub>, and find that there is no long run effect of  $x$  on  $y$  ( $F_{y,x} = 0$ ). Given that there is no long run effect, we simply conclude that there is no speed of adjustment to measure.

We note that  $\Delta X_{2t}$  has no feedback from the equilibrium correction term also in DGP<sub>2</sub>, although the  $\Delta X_{2t}$  process is no longer a pure random walk, but a more

elaborate multivariate process. In  $DGP_2$  it turns out that  $F_{\Delta X_2, \Delta X_2} = \frac{1}{3}$ , so that there is a non-zero long run effect, and hence it is natural to define and measure the various times of upcrossing in order to measure the half-life of  $\Delta X_2$  adjustment. In  $DGP_2$  one finds  $u_1 = u_{\max} = 3$ . Note that this is not an instantaneous half-life. This illustrates how multivariate systems offer much richer situations than univariate processes, exemplified here by random walks. The conclusion that, in the presence of zero error correction coefficients in the VEC, the corresponding variable does not adjust to long-run equilibrium is hence not warranted. Similarly a zero error correction coefficient does not imply that the corresponding equation or variable adjusts faster to disequilibrium.

The simple examples discussed in this section point out that the size of a long run effect and the speed of adjustment are two distinct concepts. In order to define the half-life as a measure of speed it is necessary that the corresponding long run effect size be non-zero. This implies that one should check whether the long run effect size are significant before computing point and interval estimators of the half-life. We next turn to general definitions.

### 3 Definitions

This section presents general definitions. Specifically, notation is introduced in Subsection 3.1 and the definition of IF is reported in Subsection 3.2. The multivariate definition of half-life is provided in Subsection 3.3, with remarks and examples. Finally in Subsection 3.4 these definitions are illustrated with reference to  $DGP_1$  and  $DGP_2$  introduced in Section 2.

#### 3.1 VAR

In this subsection we introduce notation for VAR processes and focus on the  $I(1)$  case, because the data for exchange rates and relative prices investigated in Section 5 displays nonstationary features (at least) of  $I(1)$  type.<sup>1</sup> Consider a  $n \times 1$  vector process  $X_t := (X_{1t}, \dots, X_{nt})'$  generated by a VAR

$$\Pi(L)X_t = \mu^* D_t^* + \epsilon_t \quad (3)$$

where  $\Pi(L) = I - \sum_{i=1}^k \Pi_i L^i$  and  $\epsilon_t$  is i.i.d.  $N(0, \Omega)$ . The deterministic component is  $\mu^* D_t^*$ , where in the empirical section we take  $D_t^* := 1$  and indicate the corresponding coefficient as  $\mu^* = \mu_1$ . Unless otherwise stated, we follow notation used in Johansen (1996) and assume  $k \geq 2$ .

The VAR can be equivalently rewritten in the form

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \mu^* D_t^* + \epsilon_t, \quad (4)$$

where  $\Gamma_i := -\sum_{j=i+1}^k \Pi_j$ . It is well known, see Johansen (1996), that process (4) generates  $I(1)$  variables with no linear trend if the following conditions hold:  $I(1)_a$ : every root  $z$  of the characteristic polynomial of  $X_t$  satisfies  $z = 1$  or  $|z| > 1$ .  $I(1)_b$ :

---

<sup>1</sup>The present approach can cover the case of stationary processes as well as of processes integrated of higher order, using appropriate definitions of the companion matrix, see Omtzigt and Paruolo (2005) for details.

$\Pi := -\Pi(1) = \alpha\beta'$ , where  $\alpha$  and  $\beta$  are  $n \times r$  matrices of full rank  $r < n$  and  $\mu_1 = \alpha\beta'_0$  with  $\beta'_0$  a  $r \times 1$  vector.  $I(1)_{-c}$ :  $\alpha'_\perp \Gamma \beta_\perp$  has full rank  $n - r$ , where  $\Gamma := -I + \sum_{i=1}^{k-1} \Gamma_i$ . We call these conditions the  $I(1)$  assumption.

Under the  $I(1)$  assumption, the VAR can be written in (many equivalent) companion forms. As in Omtzigt and Paruolo (2005), we let  $\tilde{X}_t := (\Delta X_t' : X_{t-1}'\beta : U_t')$  be the state vector, where  $U_t := (\Delta X_{t-1}' : \dots : \Delta X_{t-k+1}')'$  and  $\beta$  is a basis of the cointegration space in Assumption  $I(1)_{-b}$ . Furthermore, define  $\Gamma_1^* := \alpha\beta' + \Gamma_1$ ,  $\Phi_1 := \Gamma_2$ ,  $\Phi_2 := (\Gamma_3 : \dots : \Gamma_{k-1})$ . The associated state space representation is

$$\tilde{X}_t = A\tilde{X}_{t-1} + u_t \quad (5)$$

with  $u_t := M(\mu^* D_t^* + \epsilon_t)$ ,  $M := (I_n : 0_{n \times m+r})'$ , and

$$\tilde{X}_t := \begin{pmatrix} \Delta X_t \\ \beta' X_{t-1} \\ U_t \end{pmatrix} \begin{matrix} n \\ r \\ m \end{matrix}$$

$$A := \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} := \left( \begin{array}{cc|cc} n & r & n & m-p \\ \Gamma_1^* & \alpha & \Phi_1 & \Phi_2 \\ \beta' & I_r & & \\ \hline I_n & & & \\ & & I_{m-n} & \end{array} \right) \begin{matrix} n \\ r \\ n \\ m-n \end{matrix} \quad (6)$$

where we have reported dimensions alongside blocks of the state vector and of the companion matrix. For brevity the  $A_{22}$  block in (6) is partitioned in blocks of  $n$  and  $m - n$  rows times  $m - n$  and  $n$  columns, unlike the other blocks. Zero entries are not reported unless when needed for clarity.

Under Assumption  $I(1)_{-a}$ , the eigenvalues of  $A$  are less than 1 in modulus, and hence the companion matrix  $A$  in (6) is stable.<sup>2</sup> Note also that the companion form (6) is formulated for  $k \geq 2$ . This assumption is not restrictive from a representation point of view, because any VAR(1) can be written as VAR(2) with a zero second order matrix coefficient. Omtzigt and Paruolo (2005) discuss how the inference procedures should be modified to account for the case of a VAR(1) also in estimation and testing.

### 3.2 Impact factors and interim multipliers

In this section we report the definition of IF taken from Omtzigt and Paruolo (2005) and discuss interim multipliers, which are used in Subsection 3.3 to define upcrossings and the half-life.

Consider the companion form  $\tilde{X}_t = A\tilde{X}_{t-1} + u_t$ . Let  $\tilde{v} := \tilde{X}_t^c - \tilde{X}_t$  be a perturbation in  $\tilde{X}_t$  which induces a change  $e_i(\tilde{v}) := E(\tilde{X}_{t+i} | \tilde{X}_t^c) - E(\tilde{X}_{t+i} | \tilde{X}_t) = A^i \tilde{v}$  in the forecast function at forecast horizons  $i = 1, \dots, \ell$ .<sup>3</sup> We consider the cumulated changes up to

<sup>2</sup>We call a square matrix  $A$  stable when its eigenvalues are less than 1 in modulus.

<sup>3</sup>Here and in the following  $E(\cdot | Z)$  indicates expectation conditional on  $Z$ , where for simplicity of notation  $Z$  also indicates the conditioning value taken by the corresponding random variable.

some finite horizon  $\ell$ , and up to  $\infty$ :

$$\begin{aligned}\text{CE}(\tilde{v}, \ell) &:= \sum_{i=1}^{\ell} e_i(\tilde{v}) = \sum_{i=1}^{\ell} A^i \tilde{v} = ((I - A^{\ell+1})(I - A)^{-1} - I) \tilde{v}, \\ \text{TE}(\tilde{v}) &:= \sum_{i=1}^{\infty} e_i(\tilde{v}) = \sum_{i=1}^{\infty} A^i \tilde{v} = ((I - A)^{-1} - I) \tilde{v}.\end{aligned}$$

Because the eigenvalues of the companion matrix  $A$  are inside the unit disc, the sum in  $\text{CE}(\tilde{v}, \ell)$  converges for  $\ell \rightarrow \infty$  to  $\text{TE}(\tilde{v})$ , which is finite. The IF and the interim forecast multiplier (IFM) are defined as the sensitivity measure of TE and CE with respect to changes  $\tilde{v}$  in  $\tilde{X}_t$ :

$$F(\ell) := \frac{\partial \text{CE}(\tilde{v}, \ell)}{\partial \tilde{v}'} = \sum_{i=1}^{\ell} A^i = (I - A^{\ell+1})(I - A)^{-1} - I, \quad (7)$$

$$F := \frac{\partial \text{TE}(\tilde{v})}{\partial \tilde{v}'} = \sum_{i=1}^{\infty} A^i = (I - A)^{-1} - I. \quad (8)$$

In the following we indicate by  $F_{y,x}$  and  $F_{y,x}(\ell)$  blocks of the IF and IFM matrix in (7) and (8) associated with forecast of  $y_t$  and variation in  $x_t$ , where  $y_t$  and  $x_t$  are sets of linear combinations of  $\tilde{X}_t$  defined respectively as  $y_t := b' \tilde{X}_t$  and  $x_t := \bar{a}' \tilde{X}_t$ , for suitable choice of the selection matrices of full column rank  $a, b$ . Here  $\bar{a} := a(a'a)^{-1}$ ; in most of the paper  $y_t$  and  $x_t$  are scalars, corresponding to  $a, b$  being selection vectors. Several remarks are in order.

**Remark 1.** With two modifications, the  $F$  matrix in (8) corresponds to the cumulative impulse response (CIR) - that is, the sum of the impulse response (IR) functions over all time horizons - computed from the vector moving average representation of the VAR, see e.g. Lütkepohl (1993), Chap 2. The two modifications regard the identity matrix appearing on the right-hand-side of (8), which is motivated by the summation starting at  $i = 1$  rather than  $i = 0$ , and the absence of any term associated with  $\Omega$ . Therefore, extending the interpretation of Andrews and Chen (1994) to a multivariate context, it is possible to regard the  $F$  matrix as a measure of persistence that summarizes the information contained in all impulse response functions of the system.

**Remark 2.** Note that  $F(\ell) \rightarrow F$ , for  $\ell \rightarrow \infty$ , because  $A^{\ell+1} \rightarrow 0$  given that  $A$  is stable. Moreover, for a 1-step ahead forecast horizon,  $F(1) := A$  is the relevant IFM. The explicit expression of  $F$  in terms of the coefficients in the VEC is given in Omtzigt and Paruolo (2005) Proposition 1.

**Remark 3.** Let  $F_{\Delta X, \tilde{X}}$  be the first block of  $n$  rows of  $F$  in (8), i.e. let  $a$  and  $b$  be matrices that select  $y_t := \Delta X_t$  and  $x_t := \tilde{X}_t$ . As observed by Bedini and Mosconi (2000),  $F_{\Delta X, \tilde{X}}$  can be interpreted as the impact factor of changes  $\tilde{v} := \tilde{X}_t^c - \tilde{X}_t$  on the levels of  $X_{\infty}$ , by interchanging derivative signs and summations in the definition of  $F$ . In fact observe that  $X_{\infty} = X_t + \sum_{i=1}^{\infty} \Delta X_{t+i}$ , and

$$\begin{aligned}E(X_{\infty} - X_t | \tilde{X}_t^c) - E(X_{\infty} - X_t | \tilde{X}_t) &= \sum_{i=1}^{\infty} \left( E(\Delta X_{t+i} | \tilde{X}_t^c) - E(\Delta X_{t+i} | \tilde{X}_t) \right) \\ &= M' \sum_{i=1}^{\infty} e_i(\tilde{v}) = M' \text{TE}(\tilde{v}).\end{aligned}$$

Differentiating with respect to  $\tilde{v}$  one finds  $F_{\Delta X, \tilde{X}} = \partial(E(X_\infty - X_t | \tilde{X}_t^c) - E(X_\infty - X_t | \tilde{X}_t)) / \partial \tilde{v}'$ ; hence one can interpret  $F_{\Delta X, \tilde{X}}$  as the long run (total) multiplier on  $X_\infty - X_t$  of changes in  $\tilde{X}_t$ .

### 3.3 Half-life

In this subsection we define upcrossings and half-lives based on the concepts of IF and IFM introduced above. Consider a specific choice of scalars  $y_t := b' \tilde{X}_t$  and  $x_t := \tilde{a}' \tilde{X}_t$ , and the corresponding long run effect  $F_{y,x} = b' F a$  and interim multiplier  $F_{y,x}(\ell) = b' F(\ell) a$ . Here and in the rest of the paper  $a, b$  are selection vectors.

When  $F_{y,x} \neq 0$ , it makes sense to measure the speed at which the IFM  $F_{y,x}(\ell)$  converges to the IF  $F_{y,x}$  by counting the number of periods that  $F_{y,x}(\ell)$  takes before a given fraction  $p$  of  $F_{y,x}$  is reached, where  $0 < p < 1$  and typically  $p = \frac{1}{2}$ . Specifically, when  $F_{y,x} \neq 0$  one can define the ratio

$$f_{y,x}(\ell) := \frac{F_{y,x}(\ell)}{F_{y,x}}, \quad \ell = 1, 2, \dots \quad (9)$$

where  $f_{y,x}(\ell)$  may or may not be monotonic in  $\ell$ . In the following we omit subscripts in  $f_{y,x}(\ell)$  for readability, unless needed for clarity.

The *time of first upcrossing of level  $p$*  is defined as the minimal forecast horizon  $u_{1,p}$  at which the fraction  $f(\ell)$  surpasses  $p$ :

$$u_{1,p} := \min\{\ell : f(\ell) \geq p\}. \quad (10)$$

Observe that the level set  $\mathcal{I} := \{\ell : f(\ell) \geq p\}$  is not empty, because  $f(\ell)$  converges to 1. When  $f(\ell)$  is not monotonic,  $f(\ell)$  may first increase above  $p$ , then decrease below it and increase again, several times; see Figure 1. One may then define

$$d_{1,p} := \min\{\ell > u_{1,p} : f(\ell) < p\}$$

as the *first time of downcrossing of level  $p$* . Here the set  $\{\ell > u_{1,p} : f(\ell) < p\}$  may be empty because  $f(\ell)$  may not decrease below  $p$  once it has reached it, i.e. when it is monotonic. In this case we define  $\min \emptyset := -1$ , a conventional value which signals that there is no first downcrossing. Similarly, define recursively for  $j = 2, 3, \dots$

$$u_{j,p} := \min\{\ell > d_{j-1,p} : f(\ell) \geq p\}, \quad d_{j,p} := \min\{\ell > u_{j,p} : f(\ell) < p\}$$

as the  *$j$ -th times of upcrossing and downcrossing* respectively, where again  $\min \emptyset := -1$ . Finally one may define the *last time of upcrossing* or *downcrossing* of level  $p$  as

$$u_{\max,p} := \max_{j=1,\dots} \{u_{j,p}\}, \quad d_{\max,p} := \max_{j=1,\dots} \{d_{j,p}\},$$

where again we adopt the convention  $\max \emptyset := -1$ .

Finally we define the (ordered) vector  $\omega_p := (u_{1,p} : \dots : u_{\max,p})'$ . We define as *half-life*  $h = h_{y,x}$  any of the following choices:

HL<sub>1</sub>  $h := u_{1,\frac{1}{2}} = \min \left( \omega_{\frac{1}{2}} \right)$ ; the half-life is defined as the first time of upcrossing of level  $\frac{1}{2}$ ;

HL<sub>2</sub>  $h := u_{\max, \frac{1}{2}} = \max\left(\omega_{\frac{1}{2}}\right)$ ; the half-life is defined as the last time of upcrossing of level  $\frac{1}{2}$ ;

HL<sub>3</sub>  $h := \text{median}\left(\omega_{\frac{1}{2}}\right)$ ; the half-life is defined as the median time of upcrossing of level  $\frac{1}{2}$ .

**Remark 4.** We observe that the times of upcrossings in  $\omega_p$  are all integers, because they correspond to forecast horizons for a discrete time processes. This follows Kilian and Zha (2002) and differs from standard practise. The definitions HL<sub>1</sub> and HL<sub>2</sub> provide integers for the half-life; also HL<sub>3</sub> for odd number of elements in  $\omega_p$  delivers an integer as half-life. Any choice of mean value of the elements in  $\omega_p$  may be used in place of the median in HL<sub>3</sub>, at the price of losing the property of the half-life to return an integer.

**Remark 5.** Given the discreteness of the definition of half-life, it turns out that  $h_{y,x}$  is a discontinuous non-differentiable function of the companion matrix  $A$ , and hence the  $\delta$ -method cannot be applied to derive the asymptotic distribution of the half-life from the one of  $A$ . However, the problem of inference on the half-life may be properly addressed without violating the discrete nature of  $h_{y,x}$ , as illustrated in Section 4 below.

**Remark 6.** The concept of half-life plays a key role in the PPP debate, see e.g. Kilian and Zha (2002) for a Bayesian perspective, and Rossi (2005) for the computation of confidence sets in the presence of ‘local-to-unity’ processes. However, the majority of studies focus on the persistence of real exchange rates in a univariate perspective, precluding therefore the possibility of investigating the convergence to parity of the individual variables that comprise real exchange rates. This limitation is overcome in the definition above.

**Remark 7.** Cheung et al. (2004) use Pesaran and Shin’s (1996, 1998) GIR functions, whereas Crowder (2004) exploits structural VEC comprising the PPP equilibrium to investigate the speed of adjustment of nominal exchange rates and prices. In this paper we follow a different route. The definition of half-life given above does not rely on any structural interpretation of VEC disturbances, but uses the IF and IFM instead. More specifically,  $h_{y,x}$  is a function of the companion matrix  $A$  only, and it does not depend on other coefficients and the error covariance structure  $\Omega$ , which incorporates structural simultaneity effects.

**Remark 8.** Kilian and Zha (2002) suggest that if the half-life corresponds to more than forty years, i.e.  $h_{y,x} \geq 40fr$ , where  $fr$  is the sampling frequency of the data ( $fr := 1$  for years,  $fr := 4$  for quarters,  $fr := 12$  for months, etc.), then the half-life should be considered economically indistinguishable from  $\infty$ . We consider a similar situation in Appendix A.2, when the eigenvalues of  $A$  are greater than 1. We argue that in this case the long run effect is equal to  $\infty$ , and therefore we set the corresponding half-life to  $\infty$ .

**Remark 9.** It is sometimes claimed that the implied speed of adjustment from an error correction mechanism (VEC) is the same for all variables of the system (Morley, 2006). The definition of half-life given above depends on the choice of  $y$  and  $x$ , and there is no reason to believe that  $h_{y,x}$  corresponding to different  $y, x$  pairs should indeed be equal. This is illustrated in Subsection 3.4 and in the application reported in Section 5 below.

### 3.4 Illustration

In this subsection we illustrate the definitions introduced above on DGP<sub>1</sub> and DGP<sub>2</sub> of Section 2. Assume  $X_t := (X_{1t} : X_{2t})'$  where  $X_{1t}$  is the log exchange rate and  $X_{2t}$  is the log of relative prices; recall  $\beta' = (1 : -1)$  so that  $\beta'X_t := q_t$  is the log real exchange rate. Some long run effects (IF) of interest correspond to the following elements of the  $F$  matrix defined in (8):

$$F_{\Delta X_1, q}, \quad F_{\Delta X_2, q} \quad (11)$$

$$F_{q, \Delta X_1}, \quad F_{q, \Delta X_2}, \quad F_{q, q} \quad (12)$$

The two IFs in (11) measure the long run effect of unit variations in the PPP deviations on  $X_1$  and  $X_2$ .  $F_{\Delta X_1, q}$  and  $F_{\Delta X_2, q}$  are called by Bedini and Mosconi (2000) the ‘long run adjustment coefficients’, because they are the natural counterparts of short run error correction coefficients  $\alpha$  in the VEC form. The three IFs in (12) measure respectively the long run impact of variations in  $\Delta X_1$ , the exchange rate depreciation rate,  $\Delta X_2$  the inflation differential, and  $q$ , the real exchange rate, on  $q$  itself.

In both DGP<sub>1</sub> and DGP<sub>2</sub> one can form the state vector  $\tilde{X}_t := (\Delta X_{1t} : \Delta X_{2t} : q_{t-1})'$  as in (5) with companion matrices

$$A := \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \gamma & 0 & 0 \\ 1 & -1 & 1 \end{pmatrix}. \quad (13)$$

Using (8) we calculate the corresponding IF for DGP<sub>1</sub> and DGP<sub>2</sub>

$$\begin{aligned} F &:= \begin{pmatrix} F_{\Delta X_1, \Delta X_1} & F_{\Delta X_1, \Delta X_2} & F_{\Delta X_1, q} \\ F_{\Delta X_2, \Delta X_1} & F_{\Delta X_2, \Delta X_2} & F_{\Delta X_2, q} \\ F_{q, \Delta X_1} & F_{q, \Delta X_2} & F_{q, q} \end{pmatrix} = \begin{pmatrix} -1 & 1 & -1 \\ 0 & 0 & 0 \\ 2 & -2 & 2 \end{pmatrix} \quad (\text{DGP}_1) \quad (14) \\ &= \begin{pmatrix} -1 & \frac{4}{3} & -\frac{4}{3} \\ 0 & \frac{1}{3} & -\frac{1}{3} \\ 2 & -\frac{8}{3} & \frac{8}{3} \end{pmatrix} \quad (\text{DGP}_2). \end{aligned}$$

Applying the definition HL<sub>1</sub> of half-life we obtain the following half-lives, arranged in a similar way as the entries in (14):

$$\begin{aligned} \begin{pmatrix} h_{\Delta X_1, \Delta X_1} & h_{\Delta X_1, \Delta X_2} & h_{\Delta X_1, q} \\ h_{\Delta X_2, \Delta X_1} & h_{\Delta X_2, \Delta X_2} & h_{\Delta X_2, q} \\ h_{q, \Delta X_1} & h_{q, \Delta X_2} & h_{q, q} \end{pmatrix} &= \begin{pmatrix} 1 & 1 & 1 \\ . & . & . \\ 1 & 1 & 1 \end{pmatrix} \quad (\text{DGP}_1) \\ &= \begin{pmatrix} 1 & 2 & 2 \\ . & 3 & 3 \\ 1 & 2 & 2 \end{pmatrix} \quad (\text{DGP}_2) \end{aligned}$$

Here ‘.’ indicates an entry where the corresponding IF is 0, and one cannot apply the definitions of half-life given above. For these systems, definitions HL<sub>1</sub>, HL<sub>2</sub> and HL<sub>3</sub> all give the same half-lives displayed above. Figure 2 plots  $f_{y,x}(\ell)$  as a function of  $\ell$  for DGP<sub>1</sub> and DGP<sub>2</sub> for all pairs  $(y, x)$  with non-zero long run effect. In the figure,  $y$  and  $x$  are identified with their position in the state vector, i.e.  $y$  and  $x$  take value 1, 2, 3, where ‘1’ corresponds to  $\Delta X_1$ , ‘2’ to  $\Delta X_2$ , ‘3’ to  $q$ .

For  $DGP_1$   $f_{y,x}(\ell)$  is identical for all  $y$  and  $x$  with nonzero long run effect. The corresponding adjustment is very fast. It can hardly be concluded, however, that  $X_{2t}$  adjust faster than  $X_{1t}$ . For  $DGP_2$ , instead, one observes different half-lives. For instance,  $q_t$  adjusts faster in response to  $\Delta X_{1t}$  than in response to  $\Delta X_{2t}$ . The speed of convergence of  $X_{1t}$  and  $X_{2t}$  in response to  $q_t$  is also different, with half-lives of 2 and 3 periods respectively.

## 4 Inference

In this section we describe how likelihood-based inference on the IF and half-life can be obtained, with special reference to the I(1) cointegrated VEC with a constant term, which is the model used to investigate PPP adjustment in Section 5. In Subsection 4.1 we recall the definition of the relevant cointegrated models which are used in the empirical analysis; we also discuss point estimators for the half-life. Confidence sets for half-lives are described in Subsection 4.2.

### 4.1 Inference on long run effects

In this section we introduce the relevant cointegrated models, and we discuss likelihood-based inference on the long run effects.

We consider the I(1) models defined in Johansen (1996). They are defined as the class of VAR processes (4) where  $\Pi = \alpha\beta'$ , with  $\alpha$  and  $\beta$  matrices of dimension  $p \times r$  and all other parameters are unrestricted, with  $\Omega$  symmetric and positive definite. Among these models we concentrate on those which exclude trend-stationary behavior. In particular we consider the model called  $H_3$  in Johansen (1996), with  $D_t^* = 1$  and  $\mu_1$  unrestricted, as well as model  $H_2$  which is the submodel of  $H_3$  where  $\mu_1 = \alpha\rho_1$ , with  $\rho_1$  unrestricted.

Likelihood based inference on the cointegration rank in these models is summarized in Johansen (1996) to which we refer for details. Once inference on the cointegration rank and on the specification of deterministic component is performed, these can be fixed in subsequent analysis. Next one can test hypothesis on  $\beta$ , like  $\beta = (1 : -1)'$ . If this test does not reject, we impose  $\beta = (1 : -1)'$ . Otherwise the cointegrating vector  $\beta$  is estimated unrestrictedly. This estimator is superconsistent, so that  $\beta$  can be considered fixed in the definition of the companion matrix  $A$ ; only  $\hat{\Gamma}_1^*$ ,  $\hat{\alpha}$ ,  $\hat{\Phi}_1$ ,  $\hat{\Phi}_2$  contribute to the first order asymptotic variance of  $\hat{A}$ .

Omtzigt and Paruolo (2005) show that  $\hat{A}$  is  $\sqrt{T}$  asymptotically normal, specifically that  $T^{1/2}R' \text{vec}(\hat{A}' - A') \xrightarrow{d} N(0, V)$  as  $T \rightarrow \infty$  with  $V$  a positive definite matrix,  $R := (I_g : 0)'$  a known selection matrix with  $g := n(n+r+m)$  columns,  $\text{vec}$  indicates the column stacking operator and  $\xrightarrow{d}$  indicates convergence in distribution.<sup>4</sup>

They also show that  $\hat{F}$ , being a function of  $\hat{A}$ , inherits a  $\sqrt{T}$  asymptotically normal distribution. This permits to calculate the Wald test for the hypothesis

$$H_0 : F_{y,x} = 0, \quad (15)$$

---

<sup>4</sup>The selection matrix  $R$  reflects the fact that  $\text{vec}(\hat{A}' - A') = T^{-1/2}R\tau_T + o_P(T^{-1/2})$ , where  $\tau_T$  is an asymptotically normal random vector with asymptotically positive definite matrix, and the remainder term is  $o_P(T^{-1/2})$ .



for scalar  $y$  and  $x$ , which provides a central tool to answer Q<sub>1</sub>: ‘does  $y$  adjust to variations in  $x$ ?’. These tests are discussed in Omtzigt and Paruolo (2005), to which we refer for details.

## 4.2 Inference on the half-life

In this subsection we discuss econometric tools to answer question Q<sub>2</sub>: ‘if  $y$  adjusts to variations in  $x$ , what is the speed of adjustment?’. The answer to this question regards point and interval estimation on the half-life  $h_{y,x}$ , where the latter is defined only when (15) is false.

We observe that  $h_{y,x}$  is a function of the companion matrix, which we express as  $h_{y,x} = h_{y,x}(A)$ ; a likelihood based estimator for the half-life is obtained as  $\hat{h}_{y,x} = h_{y,x}(\hat{A})$  i.e. by substituting  $A$  with  $\hat{A}$  as an argument of the function  $h_{y,x}(\cdot)$ .  $\hat{h}_{y,x}$  is hence the likelihood-based, plug-in estimator of the half-life.

We next consider the question of defining confidence sets for  $h_{y,x}$ . We first fix some notation. Let  $\mathcal{A}$  be a confidence set (an ellipsoid) for the companion matrix  $A$ , obtained using the asymptotic normality of  $\hat{A}$ ; specifically,  $\mathcal{A} := \{A : T \text{vec}(\hat{A}' - A')' R V^{-1} R' \text{vec}(\hat{A}' - A') \leq \chi_{1-\eta}^2(g)\}$  where  $R$  and  $V$  are defined in Section 4.1, and  $\chi_{1-\eta}^2(g)$  is the  $1 - \eta$  quantile of a  $\chi^2$  distribution with  $g$  degrees of freedom.<sup>5</sup> For large samples,  $T \rightarrow \infty$ , one has  $\Pr(A \in \mathcal{A}) \rightarrow 1 - \eta$ .

The following proposition notes that the set  $\mathcal{H}_{y,x}$  of all values of the half-life  $h_{y,x}$  obtained for any choice of  $A \in \mathcal{A}$  provides a confidence set for  $h_{y,x}$ . In order to emphasize that the following proposition does not depend on convergence results, we state it for a confidence set  $\mathcal{A}$  for which  $\Pr(A \in \mathcal{A}) = 1 - \eta$ .

**Proposition 1** *Let  $\mathcal{A}$  be a confidence set for  $A$ , i.e.  $\Pr(A \in \mathcal{A}) = 1 - \eta$ . Let the set  $\mathcal{H} := \{h, h = h(A), A \in \mathcal{A}\}$  be the corresponding set of values  $h$ , where  $h$  is any measurable function of  $A$ , possibly discrete. Then*

$$\Pr(h \in \mathcal{H}) \geq \Pr(A \in \mathcal{A}) = 1 - \eta,$$

*i.e.  $\mathcal{H}$  is a confidence set for  $h$  with coverage probability at least equal to  $1 - \eta$ .*

**Proof.** Let  $h^{-1}(\mathcal{H})$  be the inverse image of  $\mathcal{H}$ . It is simple to see that  $\mathcal{A} \subset h^{-1}(\mathcal{H})$ , so that  $\Pr(h \in \mathcal{H}) \geq \Pr(A \in \mathcal{A}) = 1 - \eta$ .  $\blacksquare$

Observe that taking  $h := h_{y,x}$  as the half-life defined by any of the  $\text{HL}_i$  definitions above,  $i = 1, 2, 3$ , satisfies the hypotheses of the proposition. One can hence conclude that  $\mathcal{H}_{y,x} := \{h_{y,x}, h_{y,x} = h(A), A \in \mathcal{A}\}$  is a confidence set for  $h_{y,x}$  at least with the same confidence level of  $\{A \in \mathcal{A}\}$ . Note that  $\mathcal{H}_{y,x}$  is a discrete set of the form

$$\mathcal{H}_{y,x} = \{h_{y,x}^{\min}, h_{y,x}^{\min} + 1, \dots, h_{y,x}^{\max} - 1, h_{y,x}^{\max}\}.$$

We say that  $h_{y,x}^{\min}, h_{y,x}^{\max}$  are the bounds of the confidence set for the half-life.<sup>6</sup>

We next discuss the problem of how to calculate the bounds of the confidence set for the half-life. This is not straightforward, because the region  $\mathcal{A}$  is high dimensional,

<sup>5</sup>One could also consider other types of one-sided regions. This possibility goes beyond the scope of this paper and it is left for future research.

<sup>6</sup>Strictly speaking  $\mathcal{H}_{y,x}$  is a discrete set and not an interval.

and direct calculation of  $h(A)$  for all values of  $A \in \mathcal{A}$  is simply unfeasible. In the Appendix we motivate a procedure that performs a grid search on the boundary of the ellipsoid  $\mathcal{A}$ . This is found to be reasonably fast and effective. In the rest of this section we summarize the arguments that support this search on the boundary of the ellipsoid  $\mathcal{A}$ .<sup>7</sup>

Let the values of  $A \in \mathcal{A}$  be decomposed into  $A = \hat{A} + B$ . The discussion reported in the Appendix consists of two arguments. We first observe in Appendix A.1 that *coeteris paribus*, half-lives depend on all the eigenvalues of  $A$  and that eigenvalues closer to 1 are associated with higher half-lives. Hence in order to find the upper bound  $h_{y,x}^{\max}$  one should try to choose  $B$  in such a way as to maximally perturb the eigenvalues of  $\hat{A}$ . We note that the largest eigenvalue of  $A = \hat{A} + B \in \mathcal{A}$  may turn out to be real and larger than 1. In this case we argue in that  $h_{y,x}$  should be set equal to  $\infty$ , see Appendix A.2.

The second argument is that matrices  $A$  with possibly more extreme eigenvalues may be found at the boundary of the ellipsoid  $\mathcal{A}$ . This argument is reported in Appendix A.3. This suggests to calculate half-lives for a grid of values of  $A$  on the boundary of the ellipsoid  $\mathcal{A}$ . Details of the grid search are reported in Appendix A.4, where we also observe how the grid can be extended to cover all of  $\mathcal{A}$ .

## 5 Empirical analysis

In this section we apply the concepts and definitions introduced above to measure the speed of adjustment to PPP of nominal exchange rates and relative prices in five industrialized countries, all investigated against the U.S. dollar. As in Cheung et al. (2004), we use cointegrated VECs and no theoretical view on the process of adjustment. The data are described in Subsection 5.1, and the analysis of specification of the VEC is outlined in Subsection 5.2. We interpret the results in Subsection 5.3. Calculations were performed in PcGive 10.0, see Doornik and Hendry (2001), and Gauss, versions 4 and 6.

### 5.1 Data

We consider monthly data on consumer prices indices (CPI) and nominal exchange rates for five industrialized countries: U.K., France, Germany, Italy and Japan, abbreviated as UK, FR, GE, IT, JP. Nominal exchange rates are expressed as national currency units per 1 U.S. dollar. CPI indices are seasonally unadjusted and have base year 2000. Data are taken from the International Monetary Fund's IFS on-line database and cover the period 1973.04–1998.12 prior to the introduction of the Euro.<sup>8</sup>

We let  $X_t := (e_t : p_t)'$ , where  $e_t$  is the log of the nominal exchange rate,  $p_t := p_t^d - p_t^{\text{US}}$ ,  $p_t^d$  is the log of domestic CPI index and  $p_t^{\text{US}}$  is the log of U.S. CPI index. It is well understood that many theories expect short-lived deviations from PPP for traded goods compared to nontraded goods, see e.g. Kim (2005) for a recent investigation. Clearly, using CPI-based measures of prices as in Cheung et al. (2004) and Crowder (2004), we deliberately choose not to control for the effect of traded and nontraded goods.

---

<sup>7</sup> Alternative search methods for finding  $h_{y,x}^{\min}$  and  $h_{y,x}^{\max}$  are left for future research.

<sup>8</sup> See <http://ifs.apdi.net/imf/logon.aspx>.

Figure 3 plots relative prices  $p_t = p_t^d - p_t^{\text{US}}$ ,  $d = \text{UK, FR, GE, IT, JP}$ , with corresponding first differences. Nominal exchange rates  $e_t$  and their first differences are plotted in Figure 4. It can be noticed from simple graphical inspection that for all countries nominal exchange rates peak around 1985. Although there is not a general consensus among economists, the agreement in September 1985 among the finance ministers and central bankers of the major industrialized countries, commonly known as the Plaza Agreement, can be regarded as a watershed in the active management of exchange rates among the industrial countries. The dynamic patterns of nominal exchange rates in Figure 3 reveal the dollar's sharp fall in value during 1986-1987 and its relative stability with respect to major European countries up the end of eighties. In this paper we follow Klein et al. (1991) and assume that a shift in the policy regime towards a more active stance in managing external imbalances through policy coordination took place in the aftermath of the Plaza Agreement.

For this reason the empirical analysis of the adjustment of nominal exchange rates and relative prices to PPP is carried out both over the 1973.04-1998.12 period and 1985.09-1998.12 sub-period.

## 5.2 Specification analysis

In this section we describe the specification strategy of the bivariate VEC  $X_t := (e_t : p_t)'$  used to investigate PPP adjustment of nominal exchange rates and (relative) prices. We defer the comment on the empirical findings on half-lives to the next subsection.<sup>9</sup>

As discussed in Section 3.3, half-lives are based on the companion matrix  $A$  associated with the VEC (4). An accurate and credible estimate of  $A$  is hence vital for measuring adjustment and its speed. This is based on a data-consistent specification of the VEC. A first step in the analysis involves the selection of lag length  $k$  for the unrestricted VAR and a misspecification analysis, which includes tests for autocorrelation of residuals as well as tests for the possible presence of  $I(2)$  components in the data.<sup>10</sup> Table 1 reports results over the entire 1973.04-1998.12 period. The results in Table 2 refer to the 1985.09-1998.12 sub-period.

For each model the number of lags  $k$  was fixed by combining standard information criteria with diagnostic tests on the residuals. For all the models we ended up with a lag-length of  $k = 2$ , with insignificant residual serial correlation, albeit with deviations from the normality assumption of the errors.<sup>11</sup>

As a further test of misspecification we tested the hypothesis that there exists a single cointegrating relation,  $r = 1$ , and an  $I(2)$  component,  $n - r - s = 1$ , see e.g. Paruolo (1996). It can be noticed that except for the case of Italy over the full sample

---

<sup>9</sup>This analysis can be easily extended to trivariate models where  $X_t := (e_t : p_t^{\text{US}} : p_t^d)'$ , as in e.g. Crowder (2004) and Bacchiocchi and Fanelli (2005); yet, the analysis based on bivariate VEC allows to keep comparisons with previous studies as simple as possible, and matches the objectives of the paper. Using  $X_t := (e_t : p_t)'$  rather than  $X_t := (e_t : p_t^{\text{US}} : p_t^d)'$  incorporates an assumption of 'symmetry', i.e. that the domestic and U.S. price levels enter the cointegrating PPP relation (if any) with coefficients having the same magnitude but opposite signs.

<sup>10</sup>This misspecification analysis appears necessary in light of the results on PPP adjustment in Bacchiocchi and Fanelli (2005) obtained on a similar period. See also Osbat (2005).

<sup>11</sup>It is well known that non-normality does not pose problems to the asymptotic properties of cointegration tests, see Johansen (1996), provided the innovations have enough moments. In the case of Japan an unrestricted dummy variable for 1997.04 was also included in the VAR to account for a relatively large variation in relative prices, as well as de-meaned seasonal dummies.

period, this test leads to a clear rejection of I(2) components, pointing therefore that I(1) cointegrated VARs can be regarded as substantially reasonable description of exchange rate and relative price dynamics over the post-Bretton Woods era.

Panel A of Tables 1 and 2 reports Johansen's (1991) likelihood ratio (LR) trace tests for cointegration rank. Panel A of Table 2 also includes the estimated largest eigenvalue of the companion matrices associated with the VEC obtained after fixing the cointegration rank of the system at  $r = 1$ , and a LR test for the specification  $H_2$  (constant restricted to the cointegration relation) against  $H_3$  (unrestricted constant).

We also considered tests for cointegration rank  $r$  jointly with the choice of deterministic parts (model  $H_2$  versus  $H_3$ , see Section 4.1), which consists in the joint selection procedure described in Johansen (1996), Chapter 12. This procedure lead to the choice of models listed in Table 2 for the subsample 1985.9-1998.12. The estimates for the whole 1973.4-1998.12 sample were found to be too persistent to warrant reliable inference (see following subsection); we have hence chosen to present evidence for the  $H_2$  model only in Table 1. Differences in the results between models  $H_2$  and  $H_3$  were negligible.

For a given choice of cointegration rank  $r$  we next performed the LR test of  $\gamma = -1$  on the cointegration vector  $\beta = (1 : \gamma)'$ , in order to check if  $\beta' X_t := q_t = e_t - p_t$  is mean-reverting. In case the test rejects the hypothesis, we interpret  $\tilde{q}_t = e_t + \hat{\gamma} p_t$ , where  $\hat{\gamma}$  is the unrestricted estimate of  $\gamma$ , as a broad measure of PPP deviations. Panel B in Tables 1 and 2 summarizes the estimated cointegrating vectors  $\beta = (1 : \gamma : \beta_0)'$  for model  $H_2$ , and  $\beta = (1 : \gamma)'$  in model  $H_3$ , with the corresponding short run adjustment coefficients  $\alpha = (\alpha_e : \alpha_p)'$ . They also report LR tests for the over-identifying restriction of long run proportionality ( $\gamma = -1$ ), with the corresponding estimates of  $\alpha$  obtained under that restriction.<sup>12</sup>

The resulting dynamics of the system was next estimated and cast in the companion form. The largest eigenvalue of these companion matrices was always close to 0.9, and the remaining 2 eigenvalues were always much smaller in magnitude, with modulus less than 0.3. This finding contrasts with the largest eigenvalues estimated over the entire 1973.4-1998.12 which were all real and larger than 0.97. When the eigenvalues of the companion form are close to 1, one can expect the asymptotic theory for stable systems to behave poorly. On the contrary the 1985.9-1998.12 subsample period is characterized by much less persistence.

The estimated IFs and implied half-lives are summarized respectively in panels A and B of Tables 3 through 7. In all cases any of the three definition of half-life  $HL_1$ – $HL_3$  gave identical results. Panel A of these tables reports the estimates of the IFs (11)-(12) defined in Section 3.2; an asterisk indicates significant IFs. The corresponding half-lives, reported in panel B of Tables 3 through 7 along with 90% confidence sets, are calculated following the procedure described in the Appendix.<sup>13</sup>

We recall that the IFs  $F_{\Delta e, q}$  and  $F_{\Delta p, q}$  capture the long run effect of PPP deviations on respectively the nominal exchange rates and relative prices (see the interpretation of Section 3.2), whereas  $F_{q, \Delta e}$  and  $F_{q, \Delta p}$  measure the long run effect on PPP deviations of respectively a unit change in the exchange depreciation rate,  $\Delta e_t$ , and in the inflation differential  $\Delta p_t$ . All these IFs play a crucial role in the assessment of the

<sup>12</sup>The wrong sign of the  $\gamma$  coefficient obtained for the UK is further confirmed by estimating a trivariate VEC for  $X_t = (e_t : p_t^{\text{US}} : p_t^{\text{UK}})'$ , i.e. by relaxing the hypothesis of symmetry, which is strongly rejected. See footnote 8.

<sup>13</sup>We also computed half-lives confidence intervals by Monte Carlo and bootstrap methods ( $10^4$  replications), without improving their width. Results are available upon request.

speed of adjustment of nominal exchange rates and prices. Indeed, only to the extent that these IFs are significant, indicating a non-negligible long run size of adjustment, one can calculate the speed of adjustment to the long run equilibrium.

Finally, the IFs  $F_{q,q}$  (or  $F_{\tilde{q},\tilde{q}}$  depending on results) reported in panel B of Tables 3 through 7, measure the long run response of PPP deviations to variations in the real exchange rates itself; the implied half-lives,  $h_{q,q}$  ( $h_{\tilde{q},\tilde{q}}$ ), can be regarded as measures of the speed of PPP deviations in response to ‘composite’ variations in nominal exchange rates and (relative) prices. These half-lives can be therefore compared with the commonly reported measures of persistence of real exchange rates, usually computed through univariate techniques.

### 5.3 Empirical findings

The main empirical results of this paper can be summarized as follows.

1. The largest eigenvalue of the companion matrix  $\hat{A}$  and the resulting half-lives over the full sample 1973.04-1998.12 are in line with the literature. The LR tests for  $\beta = (1: -1: \beta_0)'$  (stationary real exchange rate) is generally rejected at the 5% level (with the exception of Japan), and only broad and highly persistent measures  $\tilde{q}_t$  of PPP deviations can be considered. Given the results over the 1985.09-1998.12 period (see below), we can ascribe this evidence to the high price differentials persistence characterizing major industrialized countries in the turbulent seventies and the first part of the eighties. As a consequence, the estimated IFs  $F_{\Delta e, \tilde{q}}$ ,  $F_{\Delta p, \tilde{q}}$ ,  $F_{\tilde{q}, \Delta e}$ ,  $F_{\tilde{q}, \Delta p}$ ,  $F_{\tilde{q}, \tilde{q}}$ , which depend on the companion matrix, are generally insignificant.<sup>14</sup> It is hence questionable to calculate the implied half-life  $h_{\Delta e, \tilde{q}}$ ,  $h_{\Delta p, \tilde{q}}$ ,  $h_{\tilde{q}, \Delta e}$ ,  $h_{\tilde{q}, \Delta p}$ ,  $h_{\tilde{q}, \tilde{q}}$ ; moreover the stable asymptotic theory cannot be expected to deliver reliable results in this case with roots very close to unity.<sup>15</sup> On the other hand, the results for the period after the Plaza Agreement seem more tenable, see below. For this reason, in the rest of the comment we only consider the subperiod 1985.09-1998.12.
2. Empirical results concerning long run effects and the speed of adjustment are more in line with PPP expectations for the subperiod 1985.09-1998.12. When imposing cointegration rank  $r = 1$ , the stationarity of  $q_t$  receives a more clear-cut support in all cases with the exception of Italy, where we consider  $\tilde{q}_t$ . Although the estimated largest eigenvalue of the companion matrices is still close to 0.9, these roots appear relatively far from 1.
3. Unlike in the full sample, many long run effects (IF) are significant over the subsample 1985.09-1998.12. In particular, the IFs  $F_{\Delta e, q}$  and  $F_{q, \Delta e}$  are strongly significant, while  $F_{\Delta p, q}$  and  $F_{q, \Delta p}$  are not significant for all country pairs (replace  $q_t$  by  $\tilde{q}_t$  for the case of Italy). These results show that the entire adjustment to PPP is buffered by nominal exchange rates. Conversely, in the long run prices do not adjust to PPP deviations, and PPP deviations do not respond to inflation differentials. Also the short run  $\alpha_p$  adjustment coefficients are not significant; hence relative prices do not adjust either in the long nor in the short

<sup>14</sup>The only exception being the UK/US case, see the left panel of Table 3.

<sup>15</sup>In this case one could resort to other approaches based on ‘local-to-unity’ asymptotic techniques as in Rossi (2005), or to the use of prior information as in Kilian and Zha (2002).

run. These results support the findings of Cheung et al. (2004) and Crowder (2004) that nominal exchange rates are affected and affect PPP disequilibrium. Overall this evidence is not at odds with e.g. Goldfajn and Valdés (1999), who point that exchange rate appreciations tend to be reversed with nominal devaluations rather than through nominal price adjustment.

4. Point estimates of the half-life of PPP deviations with respect to variations in the real exchange rate,  $h_{q,q}$  ( $h_{\tilde{q},\tilde{q}}$ ), are less than one year. The typical upper bound of 90% confidence sets for half-lives (excluding Britain and Japan) seems slightly above the 3 to 5 years consensus documented in Rogoff (1996) and Murray and Papell (2002). From an econometric point of view, large confidence sets for half-lives reflect uncertainty associated with the estimation of the system dynamics in a relatively short sample ( $T = 158$ ). Interpreting our findings along the lines of Rossi (2005), one can argue that point estimates may be reconciled with the view that short run deviations from parity are the by-product of monetary and financial shocks.
5. Point estimates of the half-life of PPP deviations in response to nominal exchange rate variations,  $h_{q,\Delta e}$ , are also within a year. Again these point estimates are coupled with large 90% confidence sets; the upper bound varies from 4.5 years to  $\infty$ .
6. The half-lives of PPP deviations with respect to variations in relative prices,  $h_{q,\Delta p}$ , are based on insignificant IFs (see point 3 above). Hence in the absence of further information we conclude that is no adjustment, and that it is questionable to compute half-lives altogether. Obviously, one might interpret this result along the lines of Cheung et al. (2004), concluding that as the dynamics of PPP deviations is mostly explained in terms of variations in nominal exchange rates, they are solely responsible for the slow adjustment, and relative prices ‘adjust quickly’. However, relative prices and the real exchange rates are unrelated both in the short and in the long run, and hence there appears to be no adjustment of relative prices. The fact that currency markets account for most of the mean-reversion of real exchange rates does not necessarily mean that the adjustment in goods markets is quick.
7. All the results documented in the points above are not invariant to possible enlargements of the information set, see e.g. Juselius (1995). Nevertheless, concluding that in bivariate VEC relative prices adjust faster than nominal exchange rates depends on how one interprets the absence of long and short run effects for relative prices. We prefer to state that prices do not adjust either in the short nor in the long run to PPP deviations.

## 6 Conclusions

In this paper we have discussed how to measure the speed of adjustment of exchange rates and prices to PPP in the context of cointegrated VECs. A multivariate version of the concept of half-life has been given. It has been pointed out that the definition of half-life of  $y$  in response to  $x$  makes sense when  $x$  exerts a long run effect on  $y$ . Long run effects can be quantified through the IFs introduced in Omtzigt and

Paruolo (2005). If there is no long run effect of  $x$  on  $y$ , i.e. the corresponding IF is zero ( $F_{y,x} = 0$ ), defining any measure of speed is questionable. This implies that one should test the significance of IFs before calculating the corresponding half-life as a measure of speed.

Our empirical results on five industrialized countries over the post-Bretton Woods period show that considering the 1985.09-1998.12 subperiod following the Plaza Agreement, point estimates of the half-life of PPP deviations can be reconciled with the predictions of sticky-price models (i.e. half-lives less than one year). Also point estimates of the half-life of PPP deviations in response to nominal exchange rate variations are less than one year. The confidence sets for these half-lives are however very wide, reflecting uncertainty in the estimation of the system dynamics.

The results for the entire 1973.4-1998.12 are more controversial, where the high inflation regimes experienced by industrialized countries in the seventies and the first part of the eighties has the effect of increasing the persistence of the system as measured by the highest unrestricted eigenvalue of the estimated VECs. As a result, IF are insignificant and the stable asymptotics cannot be expected to provide a reliable guide for inference.

According to Cheung et al. (2004) and Engel and Morley (2001), the root of the PPP puzzle may lie in the different speeds of convergence of exchange rates and prices; the PPP puzzle should be rethought to recognize the pivotal role nominal exchange rate adjustment plays in determining the PPP reversion rate. They also highlight the different role of nominal exchange rates and relative prices in the dynamics of adjustment to PPP. Our results on the 1985-1998 period do confirm the pivotal role of nominal exchange rates in the long run adjustment to PPP, supporting the insights of Engel and Morley (2001) and Cheung et al. (2004); see also Crowder (2004). Nevertheless, it can be hardly concluded that relative prices adjust to PPP faster than nominal exchange rates. Rather, VEC estimates show that relative prices do not adjust to PPP neither in the short run, nor in the long run. Moreover, PPP deviations do not respond in the long run to inflation differentials, making the calculation of the corresponding half-life questionable. This evidence does not seem to contrast the idea that a pass-through from exchange rates to prices is incomplete and/or extremely slow.

## References

- Andrews, D.W.K, Chen, H. Y., (1994), Approximately median-unbiased estimation of autoregressive models, *Journal of Business and Economic Statistics* 12, 187-204.
- Bacchiocchi, E., Fanelli, L. (2005), Testing the purchasing power parity through I(2) cointegration techniques, *Journal of Applied Econometrics* 20, 749-770.
- Bedini, C., Mosconi, R. (2000), New tools for the dynamic analysis of cointegrated VAR models, Mimeo, Department of Economics and Production, Politecnico di Milano.
- Cheung, Y.W., Lai, K. S., (2000), On the purchasing power parity puzzle, *Journal of International Economics* 52, 321-330.
- Cheung, Y.W., Lai, K. S., Bergman, M. (2004), Dissecting the PPP puzzle: the unconventional roles of nominal exchange rate and price adjustments, *Journal of International Economics* 64, 135-150.

- Crowder, W.J. (2004), The converge of nominal exchange rates and price levels to the PPP equilibrium, mimeo, Department of Economics, University of Texas at Arlington, available at <http://www.uta.edu/faculty/crowder/WCVITA.html>.
- Dornbusch, R. (1976), Expectations and exchange rate dynamics, *Journal of Political Economy* 84, 1161-1176.
- Doornik, J. A., Hendry, D. F. (2001), *Modelling Dynamic Systems Using PcGive*, Timberlake Consultants Press, London.
- Elliot, G., Pesavento, E. (2006), On the failure of PPP for bilateral exchange rates after 1973, *Journal of Money Credit and Banking*, 38, 1405-1430.
- Engel, C. and Morley, J. C. (2001), The adjustment of prices and the adjustment of the exchange rate, *NBER Working Paper* W8550.
- Goldfajn, I., Valdés, R. (1999), The aftermath of appreciations, *Quarterly Journal of Economics* 114, 229-262.
- Golub G. H., and C. F. Van Loan (1996) *Matrix computations*, 3rd edition. Baltimore: The John Hopkins University Press
- Johansen, S. (1996), *Likelihood-based inference in cointegrated Vector Auto-Regressive models*, Oxford University Press, Second revised version, Oxford.
- Juselius, K. (1995), Do purchasing power parity and uncovered interest rate parity hold in the long run?, *Journal of Econometrics* 69, 211-240.
- Horn R. A., C. R. Johnson (1985) *Matrix analysis*. Cambridge: Cambridge University Press
- Kilian, L., Zha, T. (2002), Quantifying the uncertainty about the half-life of deviations from PPP, *Journal of Applied Econometrics* 17, 107-125.
- Kim, J. (2005), Convergence rates to purchasing power parity for traded and non-traded goods: a structural error-correction model approach, *Journal of Business and Economic Statistics* 23, 76-86.
- Klein, M., Mizrach, B., Murphy, R.G. (1991), Managing the dollar: has the Plaza Agreement mattered?, *Journal of Money Credit and Banking* 23, 742-751.
- Lütkepohl, H. (1993), *Introduction to multiple time series analysis*, Springer-Verlag, New York.
- Morley, J.C. (2006), The slow adjustment of aggregate consumption to permanent income. *Journal of Money Credit and Banking*, forthcoming.
- Muirhead R.J. (1982), *Aspects of multivariate statistical theory*, John Wiley, New York.
- Murray, C.J., Papell, D.H. (2002), The purchasing power parity persistence paradigm, *Journal of International Economics* 56, 1-19.
- Mussa, M. (1982), A model of exchange rate dynamics, *Journal of Political Economy* 90, 74-104.
- Omtzigt, P., Paruolo, P. (2005), Impact factors, *Journal of Econometrics* 128, 31-68.
- Osbat, C. (2005), PPP and international price adjustment: what role for exchange rates?, mimeo, available at <http://www.ualg.pt/feua/conf/urct/prog/ps1/ps1.htm>
- Paruolo, P. (1996), On the determination of integration indices in I(2) systems, *Journal of Econometrics* 72, 313-356.
- Pesaran, M.H., Shin, Y. (1996), Cointegration and speed of convergence to equilibrium, *Journal of Econometrics* 71, 117-143.
- Pesaran, M.H., Shin, Y. (1998), Generalized impulse response analysis in linear multivariate models, *Economic Letters* 58, 17-29.
- Rogoff, K. (1996), The purchasing power parity puzzle, *Journal of Economic Literature* 34, 647-668.
- Rossi, B. (2005), Confidence sets for half-life deviations from purchasing power parity, *Journal of Business and Economic Statistics* 23, 432-442.



# A Appendix

In this Appendix we include various technical arguments concerning the calculation of bounds for the confidence set  $\mathcal{H}_{y,x}$  for the half-life  $h_{y,x}$ . The Appendix is divided into the following subsections: Subsection A.1 discusses the relation between eigenvalues of the companion matrix  $A$  and half-lives; Subsection A.2 argues that when one eigenvalue of  $A$  is greater than or equal to 1, then all corresponding half-lives should be set equal to infinity. Subsection A.3 presents arguments that suggest to inspect half-lives calculated for values of  $A$  on the boundary of the confidence ellipsoid  $\mathcal{A}$ . Finally Subsection A.4 reports how we generated a grid of values on the boundary of  $\mathcal{A}$ .

## A.1 Eigenvalues and half-life

In this subsection we analyze the relation between the eigenvalues of  $A$  and the definition of  $h_{y,x}$ . We assume that  $F_{y,x} \neq 0$  and let  $A = SJS^{-1}$  be the Jordan canonical decomposition of  $A$ . Specifically  $J := \text{diag}(J_1, \dots, J_s)$ , the  $i$ -th Jordan block  $J_i$  has dimension  $n_i$ ,  $J_i := J_i(\lambda_i) := \lambda_i I_{n_i} + N_i$  and  $N_i$  is a matrix with all zero entries and ones on the first superdiagonal; in this representation the  $\lambda_i$  may not be distinct. We also partition  $S = (S_1 : \dots : S_s)$  and  $S^{-1} = (T_1 : \dots : T_s)'$  conformably with  $J$ , so that  $A = \sum_{i=1}^s S_i J_i T_i'$ .

Let  $\text{sign}(x)$  be the signum function which equal 1 for nonnegative  $x$  and  $-1$  for negative  $x$ , where  $x$  belongs to the extended real line, i.e.  $x$  can also be  $\pm\infty$ . We assume that  $F_{y,x} \neq 0$  and let  $c := \text{sign}(F_{y,x})$ . The condition for the level set  $\mathcal{I}$ , see (10), can be written as

$$c(F_{y,x}(\ell) - pF_{y,x}) \geq 0. \quad (16)$$

For simplicity assume that  $c = 1$ ; in the opposite case  $c = -1$ , the following inequalities need to be reversed. Using (7) and (8), eq. (16) can be written as follows

$$d(\ell) := b'F((1-p)I - A^\ell)a \geq 0 \quad (17)$$

Substituting  $A = SJS^{-1}$  one finds that  $F = S((I - J)^{-1} - I)S^{-1} =: SKS^{-1}$ . Note that  $K$  is block diagonal, with blocks

$$\begin{aligned} K_i &:= (I_{n_i} - J_i)^{-1} - I_{n_i} = ((1 - \lambda_i)I_{n_i} - N_i)^{-1} - I_{n_i} = \\ &= \rho_i \sum_{j=0}^{n_i-1} \rho_i^j N_i^j - I_{n_i} = (\rho_i - 1)I + \rho_i \sum_{j=1}^{n_i-1} \rho_i^j N_i^j \end{aligned} \quad (18)$$

$i = 1, \dots, s$ , where  $\rho_i := (1 - \lambda_i)^{-1}$ ,  $\rho_i - 1 = \lambda_i(1 - \lambda_i)^{-1}$ . Here we have used the fact that  $N_i^l = 0$ , i.e.  $N_i$  is nilpotent with index  $l$ .

Hence the left-hand-side of (17) equals

$$\begin{aligned} d(\ell) &= b'S((I - J)^{-1} - I)((1-p)I - J^\ell)S^{-1}a \\ &= \sum_{i=1}^s b'_i K_i ((1-p)I_{n_i} - J_i^\ell) a_i \end{aligned} \quad (19)$$

where  $b'_i := b'S_i$ ,  $a_i := T_i'a$ .

We next wish to express the dependence of  $J_i^\ell$  on  $\lambda_i$  explicitly. If  $\lambda_i = 0$ , then  $J_i^\ell = N_i^\ell$ , which becomes 0 for  $\ell \geq n_i$ , because  $N_i$  is nilpotent with index  $n_i$ . In this

case we set  $L_i(\ell) := N_i^\ell$  in the formulas below. Consider next the case of  $\lambda_i \neq 0$ ; similarly to (18) one has, setting  $u := \min(\ell, n_i - 1)$  and using the binomial formula for matrices,

$$J_i^\ell = \lambda_i^\ell (I_{n_i} + \lambda_i^{-1} N_i)^\ell = \lambda_i^\ell \sum_{j=0}^u \binom{\ell}{j} \lambda_i^{-j} N_i^j =: \lambda_i^\ell I_{n_i} + \lambda_i^\ell L_i(\ell)$$

where  $L_i(\ell) := \sum_{j=1}^u \binom{\ell}{j} \lambda_i^{-j} N_i^j$ . When  $\ell \rightarrow \infty$ ,  $\lambda_i^\ell L_i(\ell) \rightarrow 0$  because the exponential term  $\lambda_i^\ell$  dominates. Hence for all values of  $\lambda_i$  the term  $\lambda_i^\ell L_i(\ell)$  vanishes for large  $\ell$ .

Equation (19) can be rewritten as

$$d(\ell) = \sum_{i=1}^s b_i' K_i ((1 - p - \lambda_i^\ell) I_{n_i} - \lambda_i^\ell L_i(\ell)) a_i =: \sum_{i=1}^s b_i^{*'} \varphi_i(\ell) a_i \quad (20)$$

where  $\varphi_i(\ell) := ((1 - p - \lambda_i^\ell) I_{n_i} - \lambda_i^\ell L_i(\ell))$ ,  $b_i^{*'} := b_i' K_i$ . The terms  $\varphi_i(\ell)$  depend on  $\ell$ , and  $a_i$ ,  $b_i^*$  do not depend on  $\ell$ . For fixed  $a_i$ ,  $b_i^*$ , the sign of the term  $d(\ell)$  depends on the eigenvalues  $\lambda_i$  that appear in all the terms  $\varphi_i(\ell)$ ,  $i = 1, \dots, s$ .

The diagonal entries of  $\varphi_i(1)$  are for instance negative when  $\lambda_i$  is real and greater than  $1 - p$ . Increasing  $\ell$ , the diagonal entries of  $\varphi_i(\ell)$  become then positive.  $d(\ell)$  may be negative for small  $\ell$  and then positive for large  $\ell$  if there are terms  $z_i := b_i^{*'} \varphi_i(\ell) a_i$  with this behavior which also dominate the other terms  $z_j$  in the sum,  $j \neq i$ .

For the case of diagonalizable  $A$ , i.e. when all Jordan blocks have size  $n_i = 1$ , then (20) takes the form

$$d(\ell) = \sum_{i=1}^m b_i^* a_i (1 - p - \lambda_i^\ell), \quad (21)$$

The expressions (20) and (21) show that a high half-life is associated with eigenvalues of  $A$  close to 1, and that the half-life depends on all the eigenvalues of  $A$ .

## A.2 Nonstable eigenvalues

In this subsection we discuss the case where TE diverges. This case is relevant also for the construction of the confidence set  $\mathcal{H}_{y,x}$ , because some  $A \in \mathcal{A}$  may present eigenvalues outside the unit circle. We are specifically concerned with real and positive eigenvalues greater than one; these may accrue as the result of perturbation of the greatest real eigenvalue of  $\hat{A}$ , which often turns out to be greater than 0.9. For these cases we argue that the half-life  $h_{y,x}$  should be set equal to  $\infty$ . We also argue that specific choices of  $y$ ,  $x$  cannot dissolve this problem, unless  $A$  is constructed to satisfy certain orthogonality restrictions. Except in this special cases, any choice of  $(y, x)$  will be affected by a single eigenvalue greater than 1.

We first discuss the case of divergent  $F_{y,x}(\ell)$ . Assume that  $F_{y,x}$  diverges to  $\pm\infty$ , so that  $c := \text{sign}(F_{y,x})$  is well defined. When  $c = 1$ , (16) reads  $F_{y,x}(\ell) \geq p \cdot \infty = \infty$ , which shows that  $h_{y,x}$  should be defined as  $\infty$ . When  $c = -1$ , (16) reads  $F_{y,x}(\ell) \leq p \cdot -\infty = -\infty$ , which also shows that  $h_{y,x}$  should be defined as equal to  $\infty$ . In words, it makes sense to define the half-life of an infinitely big effect as  $\infty$ .

We next discuss the effects of the choice of  $y$ ,  $x$ . Assume that  $A$  has eigenvalues outside the unit disk. Let  $A = SJS^{-1}$  be the Jordan canonical decomposition of  $A$ , and let  $J := \text{diag}(J_1, J_2)$  be a partition of the Jordan matrix into  $J_1$ , a block

containing the eigenvalues on or outside the unit disk and into  $J_2$  a block containing the eigenvalues inside the unit disk. Let also  $S := (S_1 : S_2)$ ,  $T := (T_1 : T_2)' := S^{-1}$  be a conformable partition of  $S$  and  $S^{-1}$ .

From (7) one has

$$F(\ell) = \sum_{i=1}^{\ell} A^i = S_1 C_1 T_1' + S_2 C_2 T_2'$$

where  $C_j := \sum_{i=1}^{\ell} J_j^i$ ,  $j = 1, 2$  are upper triangular matrices. Now  $C_2 := \sum_{i=1}^{\ell} J_2^i \rightarrow (I - J_2)^{-1} - I$  for  $\ell \rightarrow \infty$ , because  $J_2$  contains the stable eigenvalues, and  $C_1 := \sum_{i=1}^{\ell} J_1^i$  diverges, because it contains eigenvalues on or outside the unit disk.

For any choice of  $y$  and  $x$ , represented by the selection vectors  $b$  and  $a$ , one sees that

$$F_{y,x}(\ell) = b' S_1 C_1 T_1' a + b' S_2 C_2 T_2' a$$

which shows that  $F_{y,x}(\ell)$  diverges unless  $b' S_1 = 0'$  or  $T_1' a = 0$ , i.e. if  $b$  is orthogonal to  $\text{col}(S_1)$  or  $a$  is orthogonal to  $\text{col}(T_1)$ , where  $\text{col}(U)$  indicates the linear subspace spanned by the columns of  $U$ . These conditions are never met unless  $A$  is constructed to satisfy them, or in case  $A$  is estimated, if it has estimated under these restrictions. We conclude that if  $A$  was not constructed to satisfy these orthogonality conditions, then for *any* choice of  $(y, x)$  one has that  $F_{y,x}(\ell)$  diverges when any eigenvalue of  $A$  falls on or outside the unit disk.

### A.3 Perturbation of eigenvalues

In this subsection we provide an argument that suggests to investigate the boundary of the ellipsoid  $\mathcal{A}$  in order to find the bounds of the confidence set  $\mathcal{H}_{y,x}$  for the half-life  $h_{y,x}$ . The argument is based on results concerning perturbation of eigenvalues.

Before stating the main argument, we recall a few results. We first state Bauer-Fike's Theorem on perturbation of eigenvalues for diagonalizable matrices, see Horn and Johnson (1985), Theorem 6.3.2 p. 365, Golub and Van Loan (1996) Theorem 7.2.2 p. 321. Here  $\mathbb{C}^{m \times n}$  indicates the set of all  $m \times n$  matrices with complex entries.  $\|\cdot\|$  indicates a matrix norm and  $\|\cdot\|$  indicates the Euclidean vector norm.

**Theorem 2 (Bauer-Fike)** *Let  $\hat{A} \in \mathbb{C}^{n \times n}$  be diagonalizable with  $\hat{A} = S \Lambda S^{-1}$  and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ . Let  $A := \hat{A} + B \in \mathbb{C}^{n \times n}$  and  $\|\cdot\|$  be a submultiplicative matrix norm such that  $\|D\| = \max_{1 \leq i \leq n} |d_i|$  for all diagonal matrices  $D := \text{diag}(d_1, \dots, d_n) \in \mathbb{C}^{n \times n}$ . If  $\lambda^\dagger$  is an eigenvalue of  $A := \hat{A} + B$ , then there is some eigenvalue  $\lambda_i$  of  $\hat{A}$  for which*

$$|\lambda^\dagger - \lambda_i| \leq \kappa(S) \|B\|$$

where  $\kappa(S) := \|S\| \|S^{-1}\|$  is the condition number of  $S$  with respect to the matrix norm  $\|\cdot\|$ .

We also recall that for example  $\|B\|_2 := \max_{\|x\|=1} \|Bx\|$  is an example of a matrix norm that satisfies Bauer-Fike's Theorem, see references above. We wish also to connect the 2-norm and Frobenius' matrix norm, defined as  $\|B\|_F := \text{tr}(B^* B)$  where  $B^*$  indicates the conjugate transpose of a complex matrix; one has

$$\|B\|_2 \leq \|B\|_F \leq \sqrt{n} \|B\|_2, \quad (22)$$

see e.g. Golub and Van Loan eq. (2.3.7) p. 56. Finally recall that  $\|B\|_F = \|\text{vec}(B)\| = \|\text{vec}(B')\|$ .

We now state the argument that suggests to investigate points on the boundary of the ellipsoid  $\mathcal{A}$ . We first notice that the ellipsoid is centered around  $\hat{A}$ , which is a diagonalizable matrix with probability one. We hence restrict attention to perturbation of generic diagonalizable matrices, where Bauer-Fike's Theorem applies in the form stated above.

Let us write the generic companion matrix  $A$  in  $\mathcal{A}$  as  $A := \hat{A} + B$  and define  $\zeta := \text{vec}(\hat{A}' - A') = \text{vec}(B') = R\varphi$ ,  $\varphi := R'\zeta$ ,  $\xi := V^{-1/2}\varphi$ ,  $q^2 := \chi_{1-\eta}^2(g)$ , so that one can write  $\mathcal{A} := \{A : \|\xi\| \leq q\}$ . Note that because  $\zeta$  lies in a subspace spanned by  $R = (I_g : 0)'$ , one has  $\|\zeta\| = \|\varphi\|$ . Note also  $\varphi = V^{1/2}\xi$ , so that by definition of the 2-norm, one has

$$\|\varphi\| \leq \|V^{1/2}\|_2 \|\xi\| = \lambda_{\max}^{1/2}(V) \|\xi\|. \quad (23)$$

Here  $\lambda_{\max}(V)$  is the largest eigenvalue of the asymptotic covariance matrix  $V$ .

Let now  $c_1 := \kappa(S)$ , where  $\hat{A} = S\Lambda S^{-1}$ ; applying Bauer-Fike's Theorem one finds, using (22) and (23), that

$$\begin{aligned} |\lambda^\dagger - \lambda_i| &\leq c_1 \|B\|_2 \leq c_1 \|B\|_F = c_1 \|\zeta\| = c_1 \|\varphi\| \\ &\leq c_1 \|V^{1/2}\|_2 \|\xi\| = c_2 \|\xi\| \end{aligned} \quad (24)$$

where  $c_2 := \kappa(S) \lambda_{\max}^{1/2}(V)$ . We hence find that a bound on the radius of the disk where the eigenvalues  $\lambda^\dagger$  of  $A$  are located with respect to the ones of  $\hat{A}$  is given by  $c_2 \|\xi\|$ . Here  $c_2$  is a function of  $\hat{A}$  and of the asymptotic covariance matrix. The ellipsoid  $\mathcal{A}$  is characterized by  $\|\xi\| \leq q$ , so that the bound on the disk radius is maximal when one selects  $\|\xi\| = q$ , i.e. at the boundary of the ellipsoid.

This suggests to perform a grid search on the boundary of the ellipsoid in order to find the bound  $h_{y,x}^{\max}$  of  $\mathcal{H}_{y,x}$ . A similar argument applies for the search of  $h_{y,x}^{\min}$ . We emphasize that this is not a proof that the companion matrix  $A$  with eigenvalues most distant from  $\hat{A}$  is on the boundary of  $\mathcal{A}$ , but rather that the wider disks in (24) are found at the boundary of  $\mathcal{A}$ . In the next subsection we describe how to generate a grid over the boundary of  $\mathcal{A}$ .

## A.4 Grid on the ellipsoid boundary

In this subsection we describe how to generate a grid of values for  $A$  on the boundary of  $\mathcal{A}$ . We first recall how one can generate a uniform grid of points on the sphere; this is used in order to define point  $\xi$  with fixed length  $q$ . Subsequently, we apply  $\varphi = V^{1/2}\xi$ , in order to obtain values of  $B$ , and eventually  $A := \hat{A} + B$  in  $\mathcal{A}$ . This grid can also be calculated for various values of  $q$ , obtaining a grid for the whole of  $\mathcal{A}$ .

Consider the following transformation from polar to rectangular coordinates, see e.g. Muirhead (1982) p. 55: let  $x_1, \dots, x_g$  denote the rectangular coordinates, and let  $\rho, \theta_1, \dots, \theta_{g-1}$  denote the polar coordinates, with  $\rho > 0$ ,  $0 \leq \theta_i \leq \pi$ ,  $i = 1, \dots, g-2$ ,

$0 < \theta_{g-1} \leq 2\pi$  and

$$\begin{aligned}
x_1 &= \rho \sin \theta_1 \cdots \sin \theta_{g-3} \sin \theta_{g-2} \sin \theta_{g-1} \\
x_2 &= \rho \sin \theta_1 \cdots \sin \theta_{g-3} \sin \theta_{g-2} \cos \theta_{g-1} \\
x_3 &= \rho \sin \theta_1 \cdots \sin \theta_{g-3} \cos \theta_{g-2} \\
&\dots \\
x_{g-1} &= \rho \sin \theta_1 \cos \theta_2 \\
x_g &= \rho \cos \theta_1.
\end{aligned}$$

In order to generate a regular grid, we first note that boundary values of  $\theta_i$ ,  $i = 1, \dots, g-2$ , are special, in the sense for instance that  $\theta_1 = 0$  is associated with the single direction  $x = (0_{g-1}, 1)$  on the sphere, no matter what values the other angles take.

We hence define a grid over values  $0 < \theta_i < \pi$ ,  $i = 1, \dots, g-2$ ,  $0 < \theta_{g-1} < 2\pi$ , and later add to the grid the unit vectors  $\pm e_{i,g}$ , where  $e_{i,g}$  is the  $i$ -th column of  $I_g$ . We choose an even integer  $s$ , and let  $v := \pi/s$ . We have chosen  $s$  even, so that  $jv$  takes on the values  $\pi/2, 3\pi/2$ . Then select  $\theta_1, \dots, \theta_{g-2} \in \{jv, j = 1, \dots, s-1\}$  and  $\theta_{g-1} \in \{jv, j = 1, \dots, 2s\}$ . This gives  $c_s := 2s(s-1)^{g-2}$  points on the sphere. We then added  $2g$  unit vectors  $\pm e_{i,g}$ , obtaining a total of  $\nu_s := c_s + 2g$  points. Note that  $c_s = O(2s^{g-1})$  grows very fast in  $s$  for given  $g$ . Some values of  $\nu_s := c_s + 2g$  are given here for  $g = 6$  as in the empirical application:  $\nu_2 = 16$ ,  $\nu_4 = 660$ ,  $\nu_6 = 7512$ ,  $\nu_8 = 38428$ ,  $\nu_{10} = 131232$ ,  $\nu_{12} = 351396$ . In the empirical section we used  $s = 8$ , i.e.  $\nu_8 = 38428$  points on the grid.

# Figures and Tables

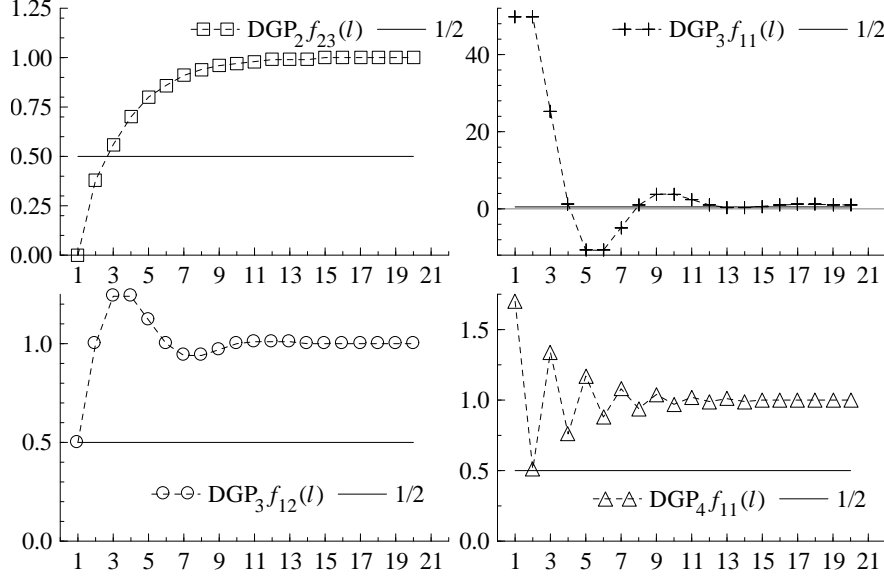


Figure 1.  $f_{y,x}(\ell)$  as a function of  $\ell$  for various choices of  $\text{DGP}_i$  ( $i = 1, 2, 3, 4$ , see Section 2) and of  $(y, x)$  pair.  $y, x$  take values 1, 2, 3, where ‘1’= $\Delta X_1$ , ‘2’= $\Delta X_2$ , ‘3’= $q$  for  $\text{DGP}_1$  and  $\text{DGP}_2$  and ‘1’= $X_1$ , ‘2’= $X_2$  for  $\text{DGP}_3$  and  $\text{DGP}_4$ . Upper-left panel: monotonic behavior with a single upcrossing. Upper-right panel: Non-monotonic behavior with more than one upcrossing. Lower-left panel: Non-monotonic behavior with a single upcrossing. Lower-right panel: Oscillating behavior, all above the threshold  $1/2$ .

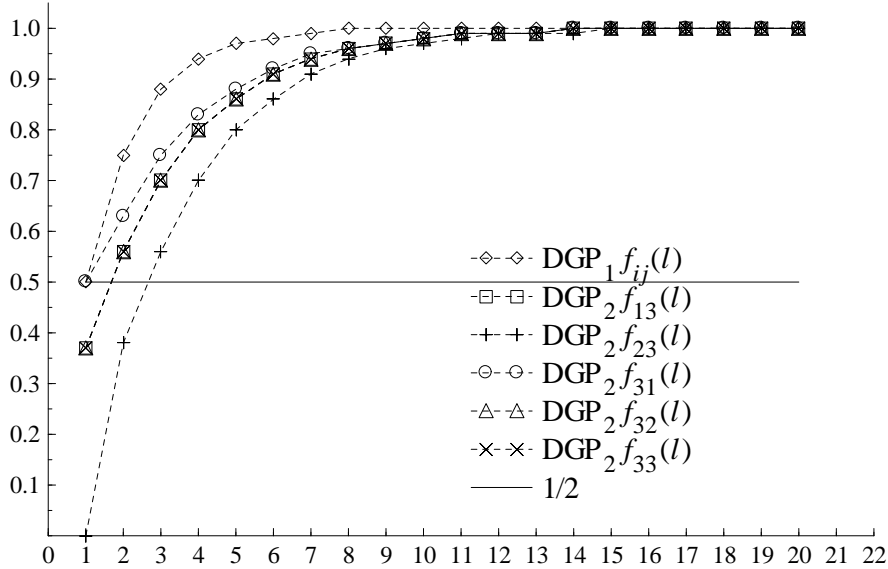


Figure 2.  $f_{y,x}(\ell)$  as a function of  $\ell$  for all possible choices of  $(y, x)$  in  $\text{DGP}_1$  and  $\text{DGP}_2$ .  $y, x$  take values 1, 2, 3, where ‘1’= $\Delta X_1$ , ‘2’= $\Delta X_2$ , ‘3’= $q$ . Only the cases with nonzero long run effect are shown. For  $\text{DGP}_1$ , the plot of  $f_{i,j}(\ell)$  is the same for  $i = 1, 2, j = 1, 2, 3$ .

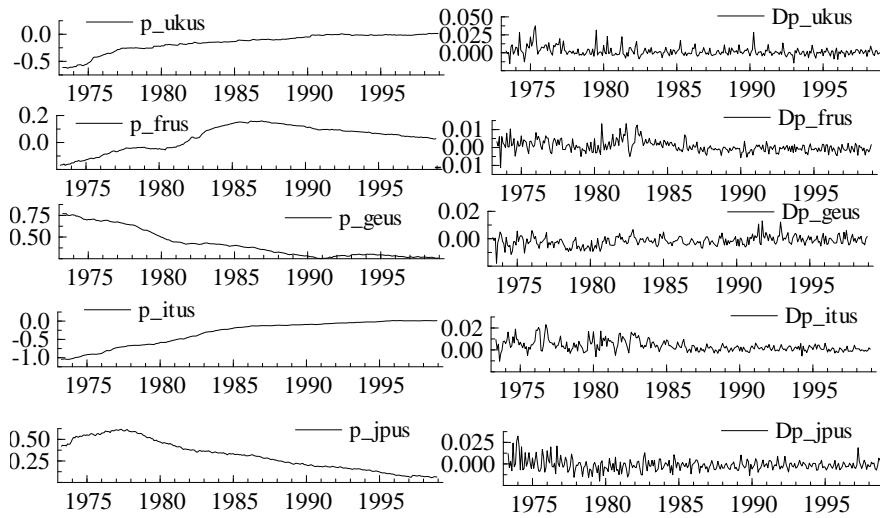


Figure 3. Relative prices  $p_t$  (left panel) and corresponding first differences (right panel).  $p_t$  is log of ratio of domestic to U.S. CPI-based prices. Monthly frequency, 1973.04-1998.12.

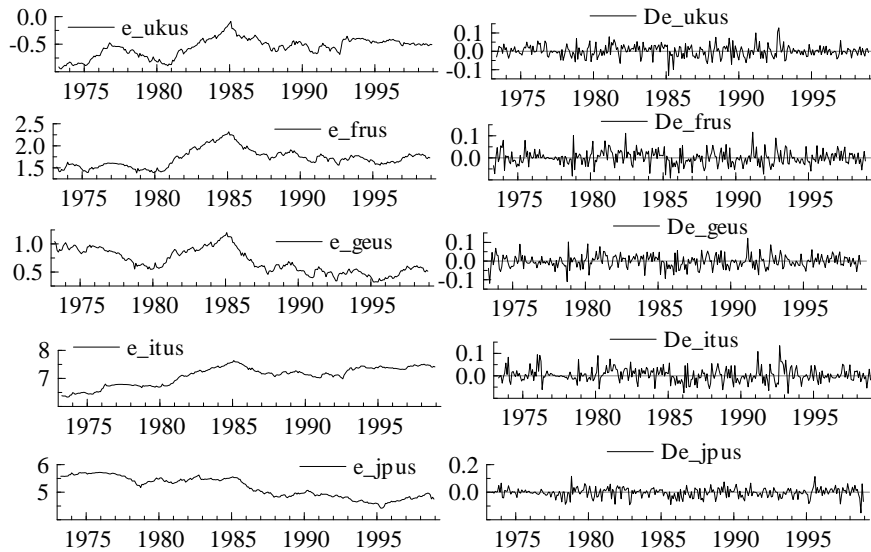


Figure 4. Nominal exchange rates  $e_t$  (left panel) and corresponding first differences (right panel).  $e_t$  is the log of domestic currency units necessary to purchase a U.S. dollar. Monthly frequency, 1973.04-1998.12.

Full Sample: 1973.04-1998.12, VEC $X_t = (e_t : p_t)'$ , $H_2$ model					
	UK/US	FR/US	GE/US.	IT/US	JP/US
A. LR trace cointegration test					
$r = 0$	41.70**	49.52**	40.82**	55.05**	28.23*
$r \leq 1$	6.15	2.9	3.55	4.44	3.04
I(2) test	194.75 [0.00]	21.49 [0.00]	27.45 [0.00]	10.67 [0.14]	19.70 [0.00]
$\lambda_{\max}(\hat{A})$	0.987	0.977	0.993	0.993	0.988
B. Estimated CI relations and over-identification tests					
$\hat{\beta} = \begin{pmatrix} 1 \\ \hat{\gamma} \\ \hat{\beta}_0 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 21.61^a \\ 0.30 \\ (4.89) \\ (1.05) \end{pmatrix}$	$\begin{pmatrix} 1 \\ -2.57 \\ -1.56 \\ (0.25) \\ (0.025) \end{pmatrix}$	$\begin{pmatrix} 1 \\ -3.25 \\ 0.096 \\ (0.65) \\ (0.28) \end{pmatrix}$	$\begin{pmatrix} 1 \\ -1.58 \\ -7.23 \\ (0.15) \\ (0.06) \end{pmatrix}$	$\begin{pmatrix} 1 \\ -1.60 \\ -4.82 \\ (0.27) \\ (0.098) \end{pmatrix}$
$\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_e \\ \hat{\alpha}_p \end{pmatrix}$	$\begin{pmatrix} -0.0004 \\ (0.0004) \\ -0.0004 \\ (0.000075) \end{pmatrix}$	$\begin{pmatrix} -0.0114 \\ (0.0119) \\ 0.0073 \\ (0.0010) \end{pmatrix}$	$\begin{pmatrix} -0.0015 \\ (0.0031) \\ 0.0019 \\ (0.0003) \end{pmatrix}$	$\begin{pmatrix} -0.0018 \\ (0.0053) \\ 0.0050 \\ (0.00067) \end{pmatrix}$	$\begin{pmatrix} 0.0022 \\ (0.0086) \\ 0.0068 \\ (0.0013) \end{pmatrix}$
LR for $\gamma = -1$	16.74 [0.00]	28.78 [0.00]	11.61 [0.00]	14.72 [0.00]	2.73 [0.10]
$\hat{\beta}_0$ when $\gamma = -1$	0.64 (0.07)	-1.59 (0.004)	-0.68 (0.08)	-7.15 (0.05)	-5.09 (0.07)
$\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_e \\ \hat{\alpha}_p \end{pmatrix}$	$\begin{pmatrix} -0.0037 \\ (0.0064) \\ 0.0057 \\ (0.0013) \end{pmatrix}$	$\begin{pmatrix} -0.0182 \\ (0.011) \\ 0.004 \\ (0.001) \end{pmatrix}$	$\begin{pmatrix} -0.0041 \\ (0.005) \\ 0.0025 \\ (0.00048) \end{pmatrix}$	$\begin{pmatrix} -0.0074 \\ (0.0074) \\ 0.0058 \\ (0.00096) \end{pmatrix}$	$\begin{pmatrix} 0.0027 \\ (0.0062) \\ 0.0046 \\ (0.001) \end{pmatrix}$

Table 1: Full sample results, 1973.04-1998.12. Panel A: LR trace test for cointegration rank and I(2) test for  $r = 1$ , and 1 I(2) component. Panel B: LR tests for  $\gamma = -1$  and adjustment coefficients. Standard errors in round brackets,  $p$ -values in square brackets.  $\lambda_{\max}(\hat{A})$ : maximal eigenvalue of the estimated companion matrix. For the JP/US model 11 de-meaned seasonal dummies were included, as well as a dummy for 1997.04.  $a$  : in the corresponding tri-variate VAR the hypothesis that the log price series have the same coefficient with opposite sign is rejected.



Subsample: 1985.09 - 1998.12, VEC $X_t = (e_t : p_t)'$					
	UK/US	FR/US	GE/US	IT/US	JP/US
A. LR trace cointegration test					
model	$H_3$	$H_3$	$H_2$	$H_3$	$H_3$
$r = 0$	13.93	15.91*	22.66*	16.79*	14.78
$r \leq 1$	3.11	0.02	6.96	1.36	5.40
I(2) test	96.67 [0.00]	81.32 [0.00]	85.70 [0.00]	77.48 [0.00]	120.6 [0.00]
$\lambda_{\max}(\hat{A})$	0.873	0.897	0.939	0.910	0.928
LR test for $H_2(1) \subset H_3(1)$	4.09 [0.04]	11.13 [0.00]	2.52 [0.11]	8.55 [0.00]	3.92 [0.048]
B. Estimated CI relations and short run adjustment coefficients					
$\hat{\beta}$	$\begin{pmatrix} 1 \\ 0.16 \\ (0.58) \end{pmatrix}$	$\begin{pmatrix} 1 \\ -0.54 \\ (0.57) \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1.71 \\ (1.39) \\ -0.95 \\ (0.41) \end{pmatrix}$	$\begin{pmatrix} 1 \\ -2.67 \\ (0.59) \end{pmatrix}$	$\begin{pmatrix} 1 \\ -0.904 \\ (0.61) \end{pmatrix}$
$\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_e \\ \hat{\alpha}_p \end{pmatrix}$	$\begin{pmatrix} -0.104 \\ (0.032) \\ -0.0025 \\ (0.0049) \end{pmatrix}$	$\begin{pmatrix} -0.0998 \\ (0.026) \\ 0.002 \\ (0.0017) \end{pmatrix}$	$\begin{pmatrix} -0.047 \\ (0.014) \\ -0.0031 \\ (0.0014) \end{pmatrix}$	$\begin{pmatrix} -0.074 \\ (0.020) \\ 0.0019 \\ (0.0014) \end{pmatrix}$	$\begin{pmatrix} -0.058 \\ (0.021) \\ 0.0019 \\ (0.0017) \end{pmatrix}$
LR test for $\gamma = -1$	1.97 [0.16]	0.62 [0.43]	2.18 [0.14]	6.81 [0.01]	0.01 [0.92]
$\hat{\alpha} = \begin{pmatrix} \hat{\alpha}_e \\ \hat{\alpha}_p \end{pmatrix}$	$\begin{pmatrix} -0.079 \\ (0.027) \\ 0.0017 \\ (0.0040) \end{pmatrix}$	$\begin{pmatrix} -0.10 \\ (0.028) \\ 0.0022 \\ (0.0018) \end{pmatrix}$	$\begin{pmatrix} -0.083 \\ (0.023) \\ -0.0027 \\ (0.0023) \end{pmatrix}$	$\begin{pmatrix} -0.067 \\ (0.023) \\ 0.0003 \\ (0.0016) \end{pmatrix}$	$\begin{pmatrix} -0.058 \\ (0.021) \\ 0.0021 \\ (0.0017) \end{pmatrix}$

Table 2: Subsample results, 1985.09-1998.12. Panel A: LR trace test for cointegration rank and I(2) test for  $r = 1$ , and 1 I(2) component. Panel B: LR tests for  $\gamma = -1$  and adjustment coefficients. Standard errors in round brackets,  $p$ -values in square brackets.  $\lambda_{\max}(\hat{A})$ : maximal eigenvalue of the estimated companion matrix. See also legend of Table 1.

UK/US									
1973.04-1998.12					1985.09-1998.12				
A. Impact Factors									
$F_{\Delta e, \tilde{q}}$	$F_{\Delta p, \tilde{q}}$	$F_{\tilde{q}, \Delta e}$	$F_{\tilde{q}, \Delta p}$	$F_{\tilde{q}, \tilde{q}}$	$F_{\Delta e, q}$	$F_{\Delta p, q}$	$F_{q, \Delta e}$	$F_{q, \Delta p}$	$F_{q, q}$
-0.0412 (0.0431)	-0.0444 (0.0020)	95.997 (33.215)	2161.67 (471.15)	78.535 (16.925)	-0.99 (0.06)	0.008 (0.06)	12.24 (4.096)	-15.31 (8.80)	10.57 (4.42)
B. Half lives (months) with 90% confidence sets									
$h_{\Delta e, \tilde{q}}$	$h_{\Delta p, \tilde{q}}$	$h_{\tilde{q}, \Delta e}$	$h_{\tilde{q}, \Delta p}$	$h_{\tilde{q}, \tilde{q}}$	$h_{\Delta e, q}$	$h_{\Delta p, q}$	$h_{q, \Delta e}$	$h_{q, \Delta p}$	$h_{q, q}$
55	55	85*	55*	54*	8*	5	8*	8	7*
2- $\infty$	32- $\infty$	31- $\infty^*$	32- $\infty^*$	31- $\infty^*$	3- $\infty^*$	1- $\infty$	3- $\infty^*$	1- $\infty$	2- $\infty^*$

Table 3: Panel A: Estimated IF  $F_{y,x}$  with corresponding standard errors (in brackets). Panel B. Estimated half-lives  $h_{y,x}$  with confidence intervals. The superscript \* indicates a significant corresponding IF.

FR/US									
1973.04-1998.12					1985.09-1998.12				
A. Impact Factors									
$F_{\Delta e, \tilde{q}}$	$F_{\Delta p, \tilde{q}}$	$F_{\tilde{q}, \Delta e}$	$F_{\tilde{q}, \Delta p}$	$F_{\tilde{q}, \tilde{q}}$	$F_{\Delta e, q}$	$F_{\Delta p, q}$	$F_{q, \Delta e}$	$F_{q, \Delta p}$	$F_{q, q}$
-0.016 (0.532)	0.395 (0.207)	43.189 (21.160)	-69.612 (48.041)	42.742 (22.206)	-0.98 (0.026)	0.01 (0.026)	10.03 (2.83)	20.41 (14.94)	9.26 (3.33)
B. Half lives (months) with 90% confidence sets									
$h_{\Delta e, \tilde{q}}$	$h_{\Delta p, \tilde{q}}$	$h_{\tilde{q}, \Delta e}$	$h_{\tilde{q}, \Delta p}$	$h_{\tilde{q}, \tilde{q}}$	$h_{\Delta e, q}$	$h_{\Delta p, q}$	$h_{q, \Delta e}$	$h_{q, \Delta p}$	$h_{q, q}$
59	30	30*	29	30	7*	6	7*	8	7*
1- $\infty$	11- $\infty$	11- $\infty^*$	1- $\infty$	11- $\infty$	3-86*	1-87	3-87*	1-131	2-87*

Table 4: Panel A: Estimated IF  $F_{y,x}$  with corresponding standard errors (in brackets). Panel B. Estimated half-lives  $h_{y,x}$  with confidence intervals. The superscript \* indicates a significant corresponding IF.

GE/US									
1973.04-1998.12					1985.09-1998.12				
A. Impact Factors									
$F_{\Delta e, \tilde{q}}$	$F_{\Delta p, \tilde{q}}$	$F_{\tilde{q}, \Delta e}$	$F_{\tilde{q}, \Delta p}$	$F_{\tilde{q}, \tilde{q}}$	$F_{\Delta e, q}$	$F_{\Delta p, q}$	$F_{z, \Delta q}$	$F_{q, \Delta p}$	$F_{q, q}$
0.244 (0.821)	0.383 (0.252)	138.818 (92.809)	-421.007 (301.474)	134.278 (91.224)	-1.07 (0.05)	-0.07 (0.05)	12.25 (3.65)	-4.33 (13.37)	11.52 (4.18)
B. Half lives (months) with 90% confidence sets									
$h_{\Delta e, \tilde{q}}$	$h_{\Delta p, \tilde{q}}$	$h_{\tilde{q}, \Delta e}$	$h_{\tilde{q}, \Delta p}$	$h_{\tilde{q}, \tilde{q}}$	$h_{\Delta e, q}$	$h_{\Delta p, q}$	$h_{q, \Delta e}$	$h_{q, \Delta p}$	$h_{q, q}$
96	94	93	93	93	8*	9	8*	5	8*
1- $\infty$	30- $\infty$	30- $\infty$	30- $\infty$	30- $\infty$	4-270*	1-270	4-270*	1-268	3-270*

Table 5: Panel A: Estimated IF  $F_{y,x}$  with corresponding standard errors (in brackets). Panel B. Estimated half-lives  $h_{y,x}$  with confidence intervals. The superscript \* indicates a significant corresponding IF.

IT/US									
1973.04-1998.12					1985.09-1998.12				
A. Impact Factors									
$F_{\Delta e, \tilde{q}}$	$F_{\Delta p, \tilde{q}}$	$F_{\tilde{q}, \Delta e}$	$F_{\tilde{q}, \Delta p}$	$F_{\tilde{q}, \tilde{q}}$	$F_{\Delta e, \tilde{q}}$	$F_{\Delta p, \tilde{q}}$	$F_{\tilde{q}, \Delta e}$	$F_{\tilde{q}, \Delta p}$	$F_{\tilde{q}, \tilde{q}}$
0.989 (1.977)	1.258 (1.251)	147.163 (142.029)	-147.753 (178.950)	136.927 (134.530)	-0.89 (0.069)	0.038 (0.026)	12.73 (3.22)	-32.33 (20.04)	11.27 (3.43)
B. Half lives (months) with 90% confidence sets									
$h_{\Delta e, \tilde{q}}$	$h_{\Delta p, \tilde{q}}$	$h_{\tilde{q}, \Delta e}$	$h_{\tilde{q}, \Delta p}$	$h_{\tilde{q}, \tilde{q}}$	$h_{\Delta e, \tilde{q}}$	$h_{\Delta p, \tilde{q}}$	$h_{\tilde{q}, \Delta e}$	$h_{\tilde{q}, \Delta p}$	$h_{\tilde{q}, \tilde{q}}$
98	96	96	95	96	8*	9	8*	8	8*
1- $\infty$	24- $\infty$	23- $\infty$	2- $\infty$	23- $\infty$	4-45*	1-79	4-45*	1-110	3-45*

Table 6: Panel A: Estimated IF  $F_{y,x}$  with corresponding standard errors (in brackets). Panel B. Estimated half-lives  $h_{y,x}$  with confidence intervals. The superscript \* indicates a significant corresponding IF.

JP/US									
1973.04-1998.12					1985.09-1998.12				
A. Impact Factors									
$F_{\Delta e,q}$	$F_{\Delta p,q}$	$F_{q,\Delta e}$	$F_{q,\Delta p}$	$F_{q,q}$	$F_{\Delta e,q}$	$F_{\Delta p,q}$	$F_{q,\Delta e}$	$F_{q,\Delta p}$	$F_{q,q}$
0.025 (1.026)	0.640 (0.641)	92.62 (93.75)	-178.04 (183.47)	81.80 (85.38)	-0.965 (0.0281)	0.0346 (0.0291)	16.173 (5.347)	-29.748 (18.635)	14.168 (5.531)
B Half-lives (months) with 90% confidence sets									
$h_{\Delta e,q}$	$h_{\Delta p,q}$	$h_{q,\Delta e}$	$h_{q,\Delta p}$	$h_{q,q}$	$h_{\Delta e,q}$	$h_{\Delta p,q}$	$h_{q,\Delta e}$	$h_{q,\Delta p}$	$h_{q,q}$
48	57	57	57	57	10*	10	10*	10	10*
1- $\infty$	12- $\infty$	12- $\infty$	12- $\infty$	12- $\infty$	4- $\infty^*$	1- $\infty$	4- $\infty^*$	1- $\infty$	4- $\infty^*$

Table 7: Panel A: Estimated IF  $F_{y,x}$  with corresponding standard errors (in brackets). Panel B. Estimated half-lives  $h_{y,x}$  with confidence intervals. The superscript \* indicates a significant corresponding IF.